

경제 현안 세미나 I

근무여건이 임금 근로자의 여행 소비행태에 미치는 영향:

소득과 근로시간에 대한 만족도를 중심으로

(계량경제학적 분석)



경제학과 201721743 노민우  
소프트웨어학과 201920746 최현수  
정치외교학과 202022035 배소희  
경제학과 202221833 이수민

# 목 차

1. 서론 .....	
1.1 연구배경 .....	
1.2 선행연구 검토 .....	
2. 여행 및 노동 현황 .....	
2.1. 여행 현황 .....	
2.2. 노동 현황 .....	
3. 연구 설계 .....	
3.1. 데이터 선택 .....	
3.2. 데이터 전처리 .....	
3.3. 데이터 기술 통계 .....	
3.4. 데이터 분포 .....	
3.5. 분석 방법 설정 .....	
4. 모형 설명 .....	
4.1 ZINB 모형 .....	
4.2 ZINB 한계효과 .....	
5. 연구 분석 .....	
5.1. OLS 분석 .....	
5.1.1. 회귀식 설정 .....	
5.1.2 회귀 분석 결과 .....	
5.1.3 시사점 .....	
5.2. ZINB 분석 .....	
5.2.1. 회귀식 설정 .....	
5.2.2. 회귀 분석 결과 .....	
5.2.3. 시사점 .....	
6. 결론 .....	
7. 부록 .....	
7.1. 포아송 분포(Poisson Distribution) .....	
7.2. 음이항 분포(Negative Binomial Distribution) .....	

# 1. 서론

## 1.1 연구배경

여행은 유람, 휴식 등을 위해 일상생활에서 벗어나 타 국가, 다른 지역으로 떠나는 일을 말한다. 여행은 노동의 피로에서 벗어나게 하되 노동 의지를 재충전시키는 기제가 된다. 또한 여행은 삶의 질 향상을 위해서도 중요하다.

국민 소득이 증가하면서 국내 여행 뿐만 아니라 해외 여행을 가는 것도 익숙한 시대가 된지 오래이다. 1980년대 들어 국내외 여건이 크게 변화하면서 외래 관광객 유치에 집중하던 흐름에서 벗어나, 늘어나는 국민의 관광 수요에 대응하고 국민 관광을 진흥하기 위해 국내 관광 활성화가 빠르게 진행되었다. 이어 2004년부터 시행된 주 5일 근무제는 관광 산업의 발전을 더욱 촉진되는 계기가 되었다. 근로자의 복지와 근로 생산성 향상을 목적으로 도입된 주 5일 근무제는 근무환경 및 근무시간이 근로자의 삶의 질과 근로 생산성에 큰 영향을 미친다는 점을 확연히 보여주었다. 국내연구<sup>1</sup>에서도 주 5일 근무제 도입에 따른 관광정책 대응 방안으로서 일과 여가의 균형 보장을 강조했다.

이렇듯 주 5일제 도입 등 임금 근로자의 복지에 대한 정부의 관심과 정책적 지원에 힘입어 국내 임금 근로자들의 평균 근무시간 만족도와 소득 만족도<sup>2</sup>가 증가하고 있다. 같은 기간, 임금 근로자들의 여행량 또한 증가하는 추세를 보였으며, 근무여건과 여행 패턴 사이에 양의(+) 상관관계가 있음을 시사했다. 이에 대해 본 연구를 통해 근무여건과 여행 패턴 사이의 관계를 계량경제학적 방법을 통하여 통계적으로 입증하고자 한다.

여행의 주요 소비 주체는 소득이 있는 국내의 모든 경제활동인구이며, 더 구체적으로는 경제활동 인구의 76.8%를 구성하는 임금 근로자가 여행의 주요 소비층이라고 할 수 있다. 그리고 임금근로자들의 여행 소비 및 다양한 행태는 시간적, 소득적 요인에 의해 가장 많이 영향을 받을 것으로 보여진다. 이러한 점을 미루어 보아, 임금근로자들의 여행 패턴과 근무여건간의 관계를 통계적 근거를 가지고 이해하는 것은 관광 산업에 있어서 경제적으로 중요한 시사점을 얻는 데에 도움을 줄 것이다.

---

<sup>1</sup> 양혜원, (2009), 주 5일제 근무제가 여가활동, 여가지출, 여가만족도에 미친 효과분석, 제 10회 한국노동패널 학술대회 논문집

<sup>2</sup> 표 1-2, 표 1-3

Unit	%										
Time	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Wage Worker Ratio	72.6	73.2	74.1	74.5	74.6	74.9	75.4	75.6	76.1	76.5	76.8

<표 1-1, 임금 근로자 비율>

### Working Time Satisfaction Trend

Time	2011	2013	2015	2017	2019	2021	2023
Mean	3.911	3.958	3.953	4.028	4.202	4.293	4.303
Frequency	13,451	13,906	1,406	14,881	14,665	14,424	14,905

\*1(Very Unsatisfied), 2(Unsatisfied), 3(Moderate), 4(Satisfied), 5(Very Satisfied)

<표 1-2, 근무시간 만족도>

### Income Satisfaction Trend

Time	2011	2013	2015	2017	2019	2021	2023
Mean	3.557	3.634	3.674	3.709	3.861	4.004	3.996
Frequency	13,503	13,927	14,745	14,883	14,670	14,419	14,886

\*1(Very Unsatisfied), 2(Unsatisfied), 3(Moderate), 4(Satisfied), 5(Very Satisfied)

<표 1-3, 소득 만족도>

## 1.2. 선행연구 검토

노동시간과 여가활동의 관계성 연구는 주 5 일제 도입의 효과성 연구에서부터 나타났다. 한국관광연구원에서 발표한 ‘주 5 일제 근무제 도입에 따른 관광정책 대응방안’ 이강욱, 김희수(2001)에서는 근로자 삶의 질 향상은 문화관광사업 등 부가가치가 높은 여가와 관련된 산업의 발전을 촉진시킬 수 있는 계기가 될 수 있을 것이라고 예측하며 근로 시간이 관광산업의 발전과 밀접한 관계가 있다는 점을 밝혔다. 국내 여행 지출에 영향을 미치는 요인 분석 연구에선 Woo Gon Kim(2002) 논문을 참고했다. 여행지출 변화를 설명하는 변수에 대해서는 가족 구성원 수, GNP, 교육 연수, 전년도 국내여행 지출, 보유 자동차 수, 환율을 고려했으며 종속변수로는 1 인당 국내 여행 지출금으로 설정했다. 분석 결과 주당 근무시간(WT)이 국내 여행 소비에 가장 많은 영향을 주는 변수로 나타났고, 가족 구성원수(FS), 교육연수(EL)이 다음으로 많은 영향을 주는 변수로 나타났다. 이와 같이 개인적 특성이 여행에 영향을 미치는 점을 확인했다. 실제 근무시간에 따른 여가 활동 참여 실태 및 만족분석(최승목, 2020)에서도 실제 주당 근무시간이 평일 및 휴일 여가 시간이 충분하고 여가 만족도 역시 높은 것으로 확인됐다. 근로여건과 여가활동이 생활만족에 미치는 영향에 대한 연구, 임금근로자를 중심으로 허종욱(2018)에서도 여가 만족도가 생활 만족도에 유의미한 영향을 미치는 것을 확인할 수 있었다. 위와 같은 선행연구 검토를 통해 근로자의 노동시간과 여가 활동, 생활 만족도에는 관계가 있다는 것을 확인할 수 있었다.

기존의 해외의 선행연구를 검토한 결과 근무여건과 여가활동의 경우 비즈니스 출장에 대한 연구가 주로 이뤄졌다. (David Holley et al. 2008; Per Gustafson, 2012) 일 가정 양립 정도에 따라 여행에 미치는 영향에 대한 개인적 특성에 대한 연구도 성별에 따른 차이(Per Gustafson, 2006)에서는 성별에 대한 차이를 보여줬지만 이외의 개인적 특성과 근무시간, 임금 만족도와 같은 변수가 여행에 미치는 영향은 설명되지 않았다. 이런 국내의 선행연구를 검토했을 때 근무환경과 임금 만족도에 따른 여행 패턴에 대한 연구는 이뤄지지 않았다고 볼 수 있다. 본 보고서는 근로시간 만족도와 소득 만족도를 핵심 변수로 설정하여 분석을 진행하였다. 이는 선행 연구들이 주로 가족 구성원 수, GNP, 교육 연수 등의 변수를 주요 변수로 설정한 것과 비교된다. 특히, 근로시간 만족도를 독립 변수로 포함하여, 주관적 만족도가 여행 행동에 미치는 영향을 상세히 분석했다는 점에서 차별화된다. 또한, 본 보고서는 데이터 전처리 과정에서 여러 해의 데이터를 통합하고, 결측치를 처리하는 등의 기술적 접근을 통해 데이터의 신뢰성을 높였다. 이는 선행 연구들이 다중공선성 문제를 해결하기 위해 주로 Ridge Regression 을 사용한 것과 비교된다.

종합적으로 본 보고서는 선행 연구들과는 다르게 근로시간과 소득 만족도가 여행 빈도와 여행 여부에 미치는 영향을 더욱 정교하게 분석하였으며, 이는 향후 근로 환경 개선 및 여행 산업 진흥을 위한 실질적인 방안을 모색하는 데 중요한 기초 자료로 활용될 것을 기대한다

## 2. 여행 및 노동 현황

### 2.1. 여행 현황

#### 국내여행

지난 수년간 대한민국의 1 인당 연평균 국내 여행 횟수는 상당한 변화를 겪어왔다. 2010년에는 1 인당 연평균 3.96 회였던 국내 여행 횟수가 지속적인 증가 추세를 보이며,

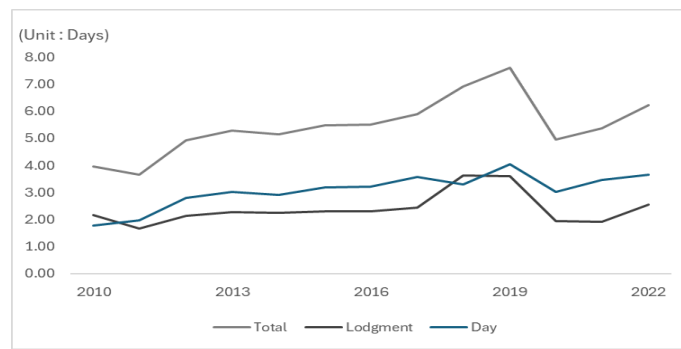
코로나 19 팬데믹 직전인 2019년에는 7.61 회로 정점을 기록하였다. 그러나 코로나 19 팬데믹의 영향으로 인해 2020년에는 4.95 회로 급격히 감소하였다. 이후 회복세를 보이며 2022년에는 6.2 회로 다시 상승하였다.

세부적으로 살펴보면, 2010년에는 1인당 숙박 여행이 연평균 2.17 회, 당일 여행이 1.79 회였다. 그러나 2019년에는 1인당 숙박 여행이 3.59 회, 당일 여행이 4.03 회로 각각 증가하였다. 팬데믹이 발생한 2020년에는 숙박 여행이 1.94 회, 당일 여행이 3.01 회로 감소하였으나, 2022년에는 숙박 여행이 2.55 회, 당일 여행이 3.66 회로 다소 회복된 모습을 보였다.

이 데이터 통해 국내 여행 횟수는 지속적으로 증가하여, 국민들의 여가 활동과 여행에 대한 관심이 높아졌음을 알 수 있다. 특히, 당일 여행의 증가폭이 두드러져, 짧은 기간 동안의 여행을 선호하는 경향이 강화되었음을 시사한다.

Unit:	Days												
Time	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Total	3.96	3.65	4.93	5.29	5.15	5.47	5.51	5.89	6.92	7.61	4.95	5.37	6.21
Lodgment	2.17	1.67	2.13	2.28	2.24	2.29	2.3	2.43	3.63	3.59	1.94	1.92	2.55
Day	1.79	1.98	2.8	3.01	2.92	3.18	3.2	3.58	3.29	4.03	3.01	3.45	3.66

<표 2-1, 1인당 국내여행 횟수>



<그래프 2-1, 1인당 연평균 국내여행 횟수>

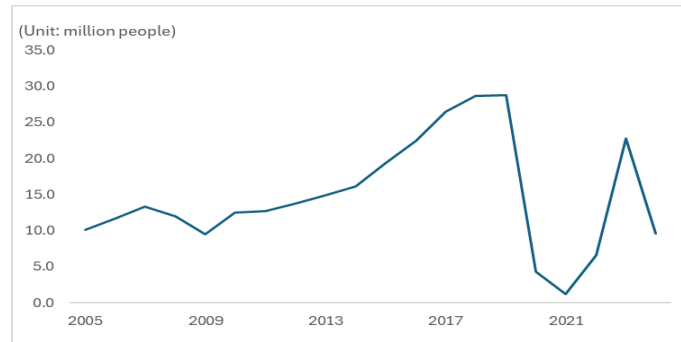
## 해외여행

지난 수년간 대한민국 국민의 총 해외여행 횟수는 상당한 변화를 겪어왔다. 2005년에는 총 1001만 명이 해외여행을 떠났으며, 이 숫자는 지속적으로 증가하여 2019년에는 2870만 명에 도달하였다. 이는 해외여행에 대한 국민들의 관심과 경제적 여건의 향상이 반영된 결과로 볼 수 있다. 그러나 코로나 19 팬데믹의 영향으로 2020년에는 해외여행 횟수가 430만 명으로 급격히 감소하였다. 이후 회복세를 보이며 2023년에는 2270만 명으로 회복되었다.

이는 국민들의 경제적 여건 향상, 저가 항공사의 증가, 여행 정보의 접근성 향상 등이 복합적 이유의 결과로 해외여행 횟수의 꾸준한 증가 추세는 국민들의 해외여행에 대한 수요가 지속적으로 증가하고 있음을 보여준다.

Unit	million people																			
Time	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
International	10.1	11.6	13.3	12.0	9.5	12.5	12.7	13.7	14.8	16.1	19.3	22.4	26.5	28.7	28.7	4.3	1.2	6.6	22.7	9.5

<표 2-2, 해외여행 총 횟수>



<그래프 2-2, 해외여행 총 횟수>

## 2.2. 노동 현황

### 평균 주당 근무시간

대한민국의 노동 환경은 지난 수년간 많은 변화를 겪어왔다. 2008 년 기준 한국의 주당 근로시간은 42.8 시간으로, 세계적으로 높은 수준에 속했다. 그러나 근로자들의 삶의 질 향상을 목표로 한 정부의 여러 정책 변화로 인해 근로시간은 점진적으로 감소하기 시작했다. 이러한 변화의 결과로, 2022년에는 주당 평균 근로시간이 36.6 시간으로 줄어들었다.

반면 2019년 평균 부여 연차일수는 15 일에서 2022년 16.6 명으로, 2019년 평균 사용연차 일수는 10.9 일에서 2022년 12.7 일로 상승하였다.

Unit:	Days			
Time	2019	2020	2021	2022
Average Granted Annual Leave	15	14.9	15.2	16.6
Average Used Annual Leave	10.9	10.7	11.6	12.7

<표 2-3. 평균 연차 사용 일수>

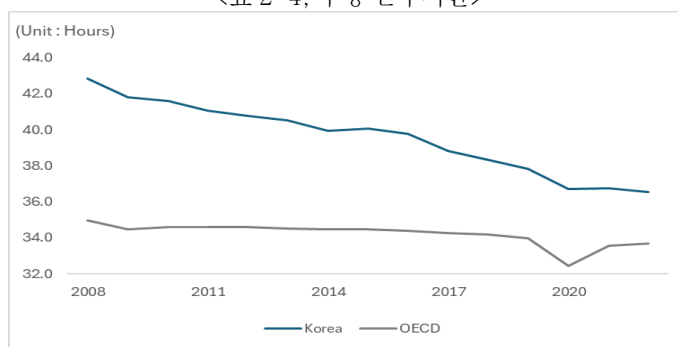
같은 기간 동안 OECD 국가들의 평균 주당 근로시간도 감소하는 추세를 보였다. 2008년에는 OECD 국가들의 평균 주당 근로시간이 35 시간이었으나, 2022년에는 33.7 시간으로

줄어들었다. 한국의 주당 근로시간은 여전히 OECD 평균보다 높은 수준을 유지하고 있지만, 감소 추세를 보이는 것은 긍정적인 변화로 평가할 수 있다.

대한민국의 평균 주당 근무시간의 변화는 노동 환경의 개선과 관련된 중요한 지표이며, 이를 통해 향후 정책 방향을 설정하고, 근로자의 삶의 질을 향상시키는 데 중요한 참고자료가 될 수 있다.

Unit:	Days														
Time	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Korea	42.8	41.8	41.6	41.1	40.8	40.5	39.9	40.1	39.8	38.8	38.3	37.8	36.7	36.7	36.6
OECD	35.0	34.5	34.6	34.6	34.5	34.5	34.5	34.5	34.4	34.2	34.0	32.4	33.5	33.5	33.7

<표 2-4, 주당 근무시간>



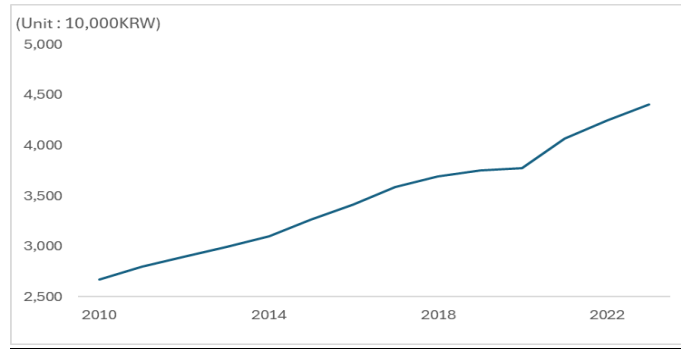
<그래프 2-3, 주당 근무시간>

## 1 인당 국민총소득

대한민국의 1 인당 국민 총소득(GNI)은 지난 수년간 상당한 성장을 보였다. 2010 년 기준으로 1 인당 국민 총소득은 2,673 만 원이었으나, 2023 년에는 4,405 만 원으로 증가하였다. 이는 경제 성장을 반영하는 지표로, 국민들의 생활 수준 향상과 직접적으로 연결된다. 특히, 일반적으로 국민들의 소득이 증가하면 여가 활동과 소비 여력이 증가하게 되며, 이는 여행 수요의 증가로 이어진다.

Unit	10,000KRW													
Time	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
GNI per capita	2,673	2,799	2,899	2,995	3,095	3,260	3,411	3,589	3,693	3,754	3,777	4,065	4,249	4,405

<표 2-5, 1 인당 GNI>



<그래프 2-4, 1인당 GNI>

### 3. 연구 설계

#### 3.1. 데이터 선택

연구 분석을 위해 노동과 근로과 관련된 Micro 범위의 데이터를 사용하기로 하였다. 따라서 본 연구에는 MDIS 마이크로데이터 통합서비스에서 제공하는 데이터 중 사회조사 데이터를 사용하였다. 사회조사 데이터는 삶의 질과 관련된 국민의 사회적 관심사와 주관적 의식에 관한 사항을 조사하여 삶의 수준과 사회적 변동을 파악하고 이를 사회개발 정책의 기초자료로 제공을 목적으로 생성된 데이터이다. 그 중 대표 데이터로 여행 횟수 및 소득 수준 등이 있기 때문에 연구를 하는 데에 사용하기 적합하다고 판단하였다.

그중 사회조사데이터의 복지/사회참여/문화와여가/소득과소비/노동 데이터를 이용하였다. 총 2011년부터 2023년까지 2년 주기로 총 7개년의 데이터를 수집하였다.

데이터 추출 방법으로는 각 년도마다 층화 2단 집락추출(Two-stage Cluster Sampling)을 이용하였다. 층화 2단 집락추출은 비슷한 성질을 갖는 2개의 이상의 동질적인 층으로 구분하고 각 층으로부터 단순 무작위 추출 방법을 통해 표본을 추출하는 방법이다. 이는

사회조사의 특성에 기인하는데 전국적으로 수도권에 인구가 밀집되어있고 전체 모집단의 목록표를 작성하지 않아도 되어 시간과 비용이 절약된다는 장점이 있다.

### 3.2. 데이터 전처리

Raw Data 자체로는 각 년도마다 조사 코드, 조사 항목 등이 일부 다르기 때문에 이를 유의하여 통일하여 통합하였다. 우선 각 연도데이터마다 연도라는 새로운 열을 만들어 해당 연도의 데이터값을 입력하였다. 두 가지의 의미를 갖게 되는데 첫 번째는 분석에 있어서 하나의 통제 변수로 사용할 수 있게 된다. 두 번째는 여러개의 연도 별 데이터를 하나의 데이터 파일로 처리하면서 표본 수를 증가시킬뿐만 아니라 각 분리된 데이터를 하나로 통합한다는데에 의미가 있다. 따라서 그 후 각 연도 데이터를 하나의 데이터파일로 통합하는 과정을 진행하였다. 통합 후의 데이터파일은 각 열별로 특정 연도의 특정 항목의 데이터만 있기 때문에 분석에 사용하고자 하는 변수들을 각 연도별로 골라내어 하나의 통합된 열로 결합하는 과정을 진행하였다. 총 43 개의 변수를 선택하였다.

성별, 만나이, 교육정도, 행정구역 등 분석을 위한 통제변수와 국내 숙박여행 횟수, 국내 당일여행 횟수, 해외여행 경험 횟수와 같은 종속변수 그리고 직업코드, 산업, 근로시간만족도, 근로 복리후생 만족도와 같은 근무 환경과 관련된 독립변수를 선택하여 분석에 선택하였다. 그 이후 각 연도 마다 코드북을 참고하여 특정 데이터가 동일한 항목을 나타나도록 조정하는 작업을 하였다.

### 3.3. 데이터 기술 통계

각 변수 별 기술 통계는 다음과 같다.

\*Summary Statistics(Key variables only)

Variables		N	Proportion
<b>Year</b>			
	2011	38,012	14.6 %
	2013	37,648	14.4 %
	2015	39,282	15.1 %
	2017	37,483	14.4 %
	2019	36,310	13.9 %
	2021	36,423	14.0 %
	2023	35,417	13.6 %
	Total	260,575	100%
<b>Sex</b>			
	1(Male)	124,675	47.8 %
	2(Female)	135,900	52.2 %
	Total	260,575	100%
<b>Age</b>			
		260,575	100%
<b>Education</b>			
	0(Uneducated)	11,156	4.3 %
	1(Elementary)	28,582	11.0 %
	2(Junior High)	32,263	12.4 %
	3(High)	84,340	32.4 %
	4(Junior College)	37,959	14.6 %
	5(Bachelor)	55,938	21.5 %
	6(Master)	8,053	3.1 %
	7(Doctoral)	2,284	0.9 %
	Total	260,575	100%
<b>Household Income</b>			
	1(Under 100k)	39,705	15.2 %
	2(100k~200K)	48,703	18.7 %
	3(200k~300k)	51,682	19.8 %
	4(300k~400k)	41,417	15.9 %
	5(400k~500k)	29,845	11.5 %
	6(500k~600k)	19,432	7.5 %
	7(Over 600K)	29,791	11.4 %
	Total	260,575	100%
<b>Household Size</b>			
	1	37,828	14.5 %
	2	77,253	29.6 %
	3	59,627	22.9 %
	4	85,867	33.0 %
	Total	260,575	100%
<b>Marital Status</b>			
	1(Unmarried)	66,068	25.4 %
	2(Married)	158,752	60.9 %
	3(bereaved)	22,969	8.8 %
	4(Divorced)	12,786	4.9 %
	Total	260,575	100%

Employment Stability							
	1(Very Unstable)		22,537				15.3 %
	2(Unstable)		59,366				40.3 %
	3(Stable)		50,151				34.0 %
	4(Very Stable)		15,298				10.4 %
	Total		147,352				100%
Subjective Income Satisfaction							
	1(Very Satisfied)		3,983				2.1 %
	2(Satisfied)		27,024				14.0%
	3(Moderate)		75,257				39.1%
	4(Unsatisfied)		61,487				31.9%
	5(Very Unsatisfied)		24,770				12.9%
	Total		192,521				100%
Subjective Working Time Satisfaction							
	1(Very Satisfied)		8,465				8.3 %
	2(Satisfied)		22,830				22.4 %
	3(Moderate)		46,355				45.5 %
	4(Unsatisfied)		16,496				16.2 %
	5(Very Unsatisfied)		6,792				6.7 %
	6(Not Sure)		928				0.9 %
	Total		101,866				100%
Variables	N	Mean	Sd	Min	Max	Skewness	Kurtosis
Count of Domestic Multi-Days Trip(Continuous)	260,575	1.1501	2.1952	0	99	5.96	94.28
Count of Domestic Single-Day Trip(Continuous)	260,575	1.9190	4.1099	0	99	5.57	58.65
Count of International Trip(Continuous)	260,575	0.2016	0.6787	0	50	10.47	354.47

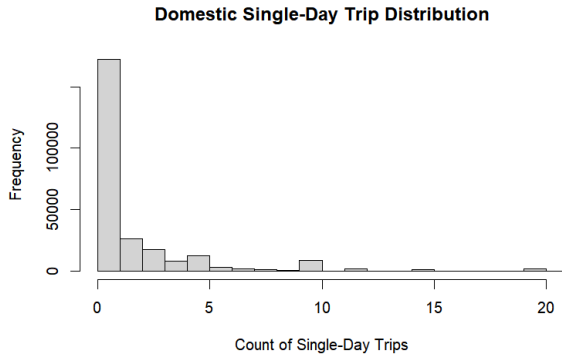
<표 3-1. 데이터 기술 통계>

종속변수로 사용되는 여행 횟수의 경우 매우 특이한 분포를 갖는다. 우선 왜도(skewness)값을 보면 각각 5.96, 5.57, 10.47로 오른쪽으로 꼬리가 긴 분포를 갖는다. 또한 median 값이 0이기 때문에 절반 이상의 사람들이 여행을 가지 않았음에 응답했음을 알 수 있다. 이는 그래프를 보면 확연히 0 값에 몰려있는 분포를 확인할 수 있다. 이는 곧 영과잉 분포로서 분석할 수 있는 근거를 갖게 된다.

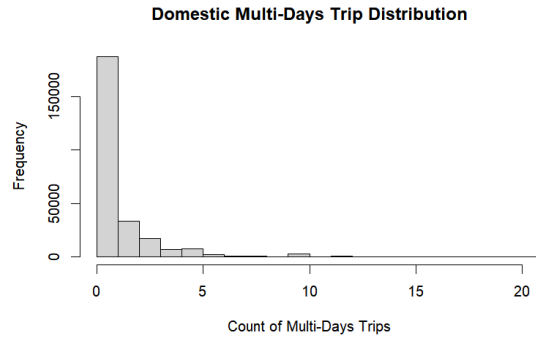
Working Time Satisfaction 변수는 모두 임금근로자를 대상으로 조사한 데이터이기 때문에 총 101,866 명의 임금근로자가 있음을 파악할 수 있다.

### 3.4. 데이터 분포

종속변수로 사용할 여행 횟수의 경우 0 값에 집중되어 있다. 이를 그래프로 나타내보면 다음과 같다.

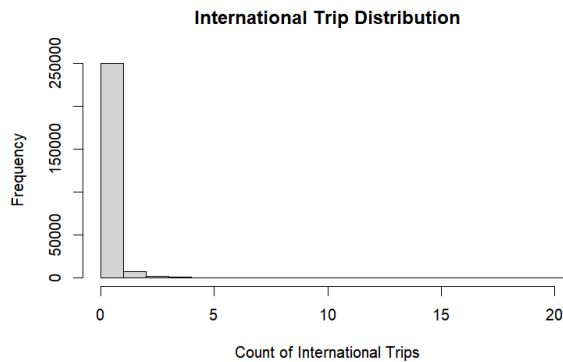


<그림 3-1. 국내 당일여행 분포도>



<그림 3-2. 국내 숙박여행 분포도>

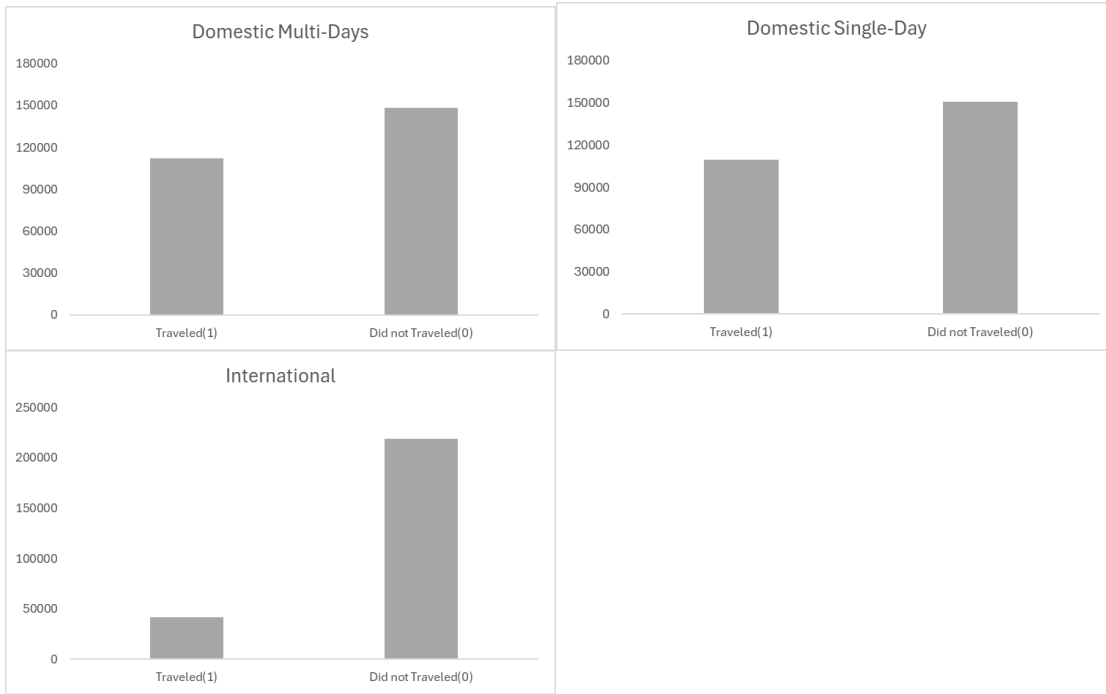
그림과 같이 국내여행의 경우 0에 상당히 많은 분포를 가진다. 특히, 당일여행보다 숙박여행이 더욱 영과잉이라고 할 수 있다.



<그림 3-3. 해외여행 분포도>

해외여행의 경우에는 국내여행보다 더 한 영과잉 분포를 띄고 있다. 종속변수를 numeric 이 아닌 binary 로 여행 유무로 보게 된다면 영에 대한 비율을 확인할 수 있다.

그래프를 통해 해외여행이 국내여행보다 더욱 더 영과잉인 상태임을 알 수 있다.



영과잉을 여부 분포도로 파악하면 확연히 파악할 수 있다. 전반적으로 각 여행 유형별로 간 사람보다 가지 않은 사람의 비율이 많음을 알 수 있다. 특히 해외여행의 경우 여행을 다녀온 사람보다 여행을 다녀오지 않은 사람의 비중이 다른 여행 유형보다 많이 차이가 남을 알 수 있다.

### 3.5. 분석 방법 설정

우선 각 근로와 관련된 데이터들이 여행 횟수 데이터에 어떤 관계성이 나타날지를 구하고자 회귀분석을 통해 본 연구의 주제인 근무시간에 따른 개인의 특성 차이에 따른 국내 여행과 해외여행 여부에 영향을 미치는지를 분석하고자 한다.

가장 접근하기 쉬운 회귀 분석 방법인 선형 회귀 분석을 통해 여행 여부 및 횟수와 근로 여건에 따른 분석을 우선적으로 진행한다.

다음으로 데이터의 분포를 고려한 ZINB 회귀 분석을 통해 여행 횟수와 근로 여건에 따른 분석을 진행할 것이다.

## 4. 모형 설명

OLS(Ordinary Least Square) 추정 방법 이외에도 데이터 분포를 고려한 다른 추정방법을 이용해 볼 수 있다. 종속변수의 경우 0 에 표본이 집중되어있는 영과잉 분포를 나타내고 있으므로 영과잉 모델을 사용하여 회귀 추정 방법을 사용할 수 있다.

영과잉 모델은 혼합 분포로서 0 만을 생성하는 분포와 Count 를 나타내는 특정 분포의 혼합 형태로 나타난다. Count 를 나타내는 분포는 정규 분포(Normal Distribution), 포아송 분포(Poisson Distribution)등 다양한 분포 중 하나로 나타날 수 있다. 물론 Count 를 나타내는 분포 또한 0 을 생성할 수 있다.

일반적으로 Count 를 나타내는 분포는 포아송 분포를 많이 사용하지만 이는 평균과 분산이 같다는 분포특성 때문에 표본이 과분산 분포를 나타낼 때 사용하기에는 적절하지 않다. 따라서 이 연구에서는 Count 분포로 음이항 분포(Negative Binomial Distribution)를 사용하였다. 이는 포아송 분포보다는 덜 제약적이며 과분산 분포를 표현하기에 적절하다.

또한, ZINB 모델을 사용함으로써 Count 에 대한 선형식과 Inflation 에 대한 선형식을 한번에 추정할 수 있는 장점이 있다. 이는 OLS 만을 사용해서는 한번에 추정할 수 없기 때문에 본 연구에 적합하다고 생각하여 사용하였다.

### 4.1 ZINB 모델

Count 분포를 음이항 분포로 한 영과잉 음이항 회귀 분석(Zero-Inflated Negative Binomial Model)을 사용할 수 있다. 우선 크게 두 가지 단계로 나눌 수 있다. 첫 번째는 표본값이 0 만을 생성하는 분포를 따르는지 Count 를 나타내는 분포를 따르는지 확률적으로 접근하여 나타낼 수 있다. 즉, 0 만을 생성하는 분포를 따를 확률  $\pi_i$  는 다음과 같이 표현된다.

$$\pi_i = \Pr(i \text{ belongs to always zero group.} \mid \text{covariates}) \quad (1)$$

$\pi_i$ 를 통한 특정 Count  $y_i$  는 다음과 같이 표현된다

$$y_i^{count} \sim \begin{cases} 0 & \text{with probability } \pi_i \\ NB(\mu_i, \theta) & \text{with probability } 1 - \pi_i \end{cases} \quad (2)$$

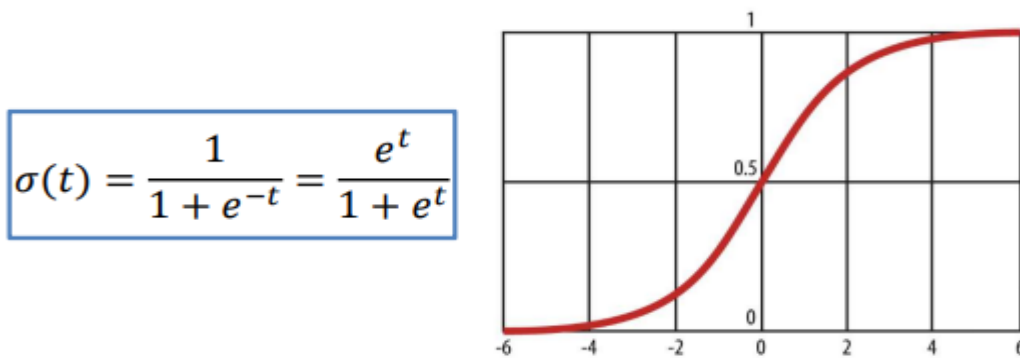
0 만을 생성하는 분포를 따를 확률  $\pi_i$ 를 특정 요인들에 의한 함수로 나타내고 싶으면

$$\pi_i = \Lambda(z_i'\gamma) = \frac{1}{1 + \exp(z_i'\gamma)} \quad (3)$$

으로 표현할 수 있다. 이 때 Logistic Function 을 활용하여 특정 요인들로 이루어진 Linear Equation 을 0 과 1 사이의 값으로 나타낼 수 있다.  $z_i'\gamma$ 에서  $z_i'$ 은 0 만을 생성하는 분포를 따르게 하는 covariates 의 vector 이며  $\gamma$ 는 각각 covariates 에 대응하는 계수값이다. Count 를 나타내는 분포에서

$$\mu_i = \exp(x_i'\beta) \quad (4)$$

로 나타나는데  $x_i'$ 는 Count 를 나타내는 분포에서 Count 를 결정하는 covariates 의 vector 이며  $\beta$ 는 각각 covariates 에 대응하는 계수값이다.  $\mu_i$ 는 관측값 i에서의 평균값을 나타내는데 일반적인 linear equation 으로는 음수 값이 나올 수도 있으므로 지수를 활용한다.



[그림 4-1. Logistic Function(Sigmoid Function)]

OLS에서는 잔차의 제곱의 합을 최소화하여 각 회귀식에서의 계수값을 구하지만, ZINB는 확률로서 나타내기 때문에 최대우도법을 통하여 계수값을 구한다.

$$L(\theta|x) = P(x|\theta) = \sum_{i=1}^n \log P(x_i|\theta)$$

최대우도법(Maximum Likelihood Estimation)이란 어떤 확률변수에서 표집한 값들을 토대로 그 확률변수의 모수를 구하는 방법이다. 조건부 확률의 곱으로 표현이 되는데 확률은 0 과 1 사이의 값을 가지며, 따라서 곱셈을 하면 할수록 그 값이 너무 작아지게 되어 회귀식의

계수를 구하기 어려워지는 문제가 발생한다 따라서 이를 자연로그를 통해 값을 증폭시키는 효과를 기대한다. ZINB 에서의 최대우도법의 식은 다음과 같다.

$$L(\beta, \gamma, \theta) = \prod_{i=1}^n [\pi_i 1(y_i = 0) + (1 - \pi_i) f(y_i; \mu_i, \theta)] \quad (5)$$

여기서 표현되는  $f(y_i; \mu_i, \theta)$ 는 음이항분포의 확률질량함수(Probability Mass Function)로 나타내며 그 식은 다음과 같다.

$$f(y_i; \mu_i, \theta) = \frac{\Gamma(y_i + \theta^{-1})}{\Gamma(y_i + 1)\Gamma(\theta^{-1})} \left( \frac{\theta^{-1}}{\theta^{-1} + \mu_i} \right)^{\theta^{-1}} \left( \frac{\mu_i}{\theta^{-1} + \mu_i} \right)^{y_i} \quad (6)$$

여기서 만약  $y_i$ 가 0 이 된다면 즉,  $f(0; \mu_i, \theta)$ 은 다음과 같이 간단하게 표현될 수 있다.

$$f(0; \mu_i, \theta) = \left( \frac{\theta^{-1}}{\theta^{-1} + \mu_i} \right)^{\theta^{-1}} \quad (7)$$

음이항 분포 Equation 에서 사용되는 감마 함수  $\Gamma(\cdot)$ 은 실수 전체의 범위에서의 팩토리얼 함수로서 나타난다. 임의의 양의 정수  $k$ 에 대해서 다음 식이 성립한다.

$$\Gamma(k) = (k - 1)! \quad (8)$$

따라서 Log Likelihood 로 나타낸 최대우도법에 대한 식은 다음과 같다.

$$\ell(\beta, \gamma, \theta) = \sum_{i=1}^n \log [\pi_i 1(y_i = 0) + (1 - \pi_i) f(y_i; \mu_i, \theta)] \quad (9)$$

위 식은 다음 식으로도 변형하여 나타낼 수 있다.

$$\begin{aligned} \ell(\beta, \gamma, \theta) &= \sum_{i=1}^n [1(y_i = 0) \ln(\pi_i + (1 - \pi_i) f(0; \mu_i, \theta)) + 1(y_i > 0) \ln((1 - \pi_i) f(y_i; \mu_i, \theta))] \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n [1(y_i = 0) \ln \left( \Lambda(z_i' \gamma) + (1 - \Lambda(z_i' \gamma)) \left( \frac{\theta^{-1}}{\theta^{-1} + \mu_i} \right)^{\theta^{-1}} \right) \\
&\quad + 1(y_i \\
&> 0) \ln \left( (1 - \Lambda(z_i' \gamma)) \frac{\Gamma(y_i + \theta^{-1})}{\Gamma(y_i + 1) \Gamma(\theta^{-1})} \left( \frac{\theta^{-1}}{\theta^{-1} + \mu_i} \right)^{\theta^{-1}} \left( \frac{\mu_i}{\theta^{-1} + \mu_i} \right)^{y_i} \right)]
\end{aligned}$$

## 4.2 ZINB 한계효과

ZINB 모델에서도 Marginal Effect 를 구하여 특정 변수  $x_k$ 가 변화 할때의 Count 기댓값  $E[Y]$  변화를 구할 수 있다. Marginal Effect 를 유도하는 식은 다음과 같다.

$$E[Y] = (1 - \pi)\mu$$

$$\pi = \Lambda(z\gamma)$$

$$\mu = e^{x\beta}$$

$$\frac{\partial \pi}{\partial x_k} = \frac{\partial \Lambda(z\gamma)}{\partial x_k} = \Lambda(z\gamma)(1 - \Lambda(z\gamma))\gamma_k$$

$$\frac{\partial \mu}{\partial x_k} = \frac{\partial e^{x\beta}}{\partial x_k} = \mu\beta_k$$

$$\frac{\partial E[Y]}{\partial x_k} = \frac{\partial}{\partial x_k} ((1 - \pi)\mu) = (1 - \pi) \frac{\partial \mu}{\partial x_k} - \mu \frac{\partial \pi}{\partial x_k}$$

$$= (1 - \pi)\mu\beta_k - \mu\Lambda(z\gamma)(1 - \Lambda(z\gamma))\gamma_k$$

$$= (1 - \pi)\mu\beta_k - \mu\pi(1 - \pi)\gamma_k$$

$$= \mu((1 - \pi)\beta_k - \pi(1 - \pi)\gamma_k)$$

이는 곧 변수  $x_k$ 가 한 단위 변화할 때  $(1 - \pi)\mu\beta_k$ 는 Count 분포에서의 Count 기댓값이 변화한다는 것을 나타내고  $-\mu\Lambda(z\gamma)(1 - \Lambda(z\gamma))\gamma_k$ 는 0 만을 생성하는 분포일 확률에 영향을 받아 Count 의 기댓값이 변화한다는 것을 나타낸다. 즉, Count Component 와 Zero-Inflation Component 에서 각각 영향을 받아  $E[Y]$ 값이 변화함을 뜻한다.

## 5. 연구 분석

### 5.1. OLS 분석

#### 5.1.1. 회귀식 설정

---

##### OLS MODEL

---

$$Y = \beta_0 + \beta_1 \text{year} + \beta_2 \text{sex} + \beta_3 \text{age} + \beta_4 \text{education} + \beta_5 \text{household income} + \beta_6 \text{household size} + \beta_7 \text{city} + \beta_8 \text{rural} + \beta_9 \text{marital status} + \beta_{10} \text{disabled} + \beta_{11} \text{employment status} + \beta_{12} \text{industry} + \beta_{13} \text{occupation} + \beta_{14} \text{employment stability} + \beta_{15} \text{subjective wage satisfaction} + \beta_{16} \text{subjective working time satisfaction} + \varepsilon$$

---

추정식에 포함된 독립변수들은 순서대로 다음과 같다.

독립 변수

- 성별: 남성(baseline), 여성
- 나이: 25 세~65 세 (연속형 변수)
- 교육수준:
  - 무학
  - 초졸
  - 중졸
  - 고졸 (baseline)
  - 2년제 대졸
  - 4년제 대졸
  - 석사 졸업
  - 박사 졸업
- 가구소득:
  - 100 만원 미만 (baseline)
  - 100~200 만원 미만
  - 200~300 만원 미만
  - 300~400 만원 미만
  - 400~500 만원 미만
  - 500~600 만원 미만
  - 600 만원 이상
- 가구원 수:
  - 1 인 (baseline)
  - 2 인
  - 3 인
  - 4 인 이상
- 도시: 국내 특별시 포함 17 개 행정구역 중 하나 (baseline: 서울)
- 시골/도시 여부: 도시 (baseline), 시골
- 결혼여부:
  - 미혼 (baseline)
  - 배우자 있음
  - 사별
  - 이혼

- 장애여부:
  - 장애 있음 (baseline)
  - 없음
- 고용 상태:
  - 상용근로자 (baseline)
  - 임시 근로자
- 산업 종류:
  - 농업, 임업 및 어업 (baseline)
  - 광업
  - 제조업
  - 전기, 가스, 증기 및 수도사업 등 총 21 개 산업군
- 직업 종류:
  - 관리자 (baseline)
  - 전문가 및 관련 종사자
  - 사무 종사자 등 총 10 개 직업군
- 고용안정성: 고용 안정의 위험성을 느끼는 정도
  - 매우 많이 느낀다 (baseline)
  - 약간 느낀다
  - 별로 느끼지 못한다
  - 전혀 느끼지 못한다
- 주관적 소득 만족도:
  - 매우 만족한다
  - 약간 만족한다
  - 보통이다 (baseline)
  - 약간 불만족한다
  - 매우 불만족한다
- 근로시간 만족도:
  - 매우 만족한다
  - 약간 만족한다
  - 보통이다 (baseline)
  - 약간 불만족한다
  - 매우 불만족한다
  - 모르겠다

나이를 제외하고는 모두 더미변수로서 추정식에 포함했다. 주된 관심사인 주관적 소득 만족도와 근로시간 만족도를 핵심변수로서 포함하였으며 나머지 변수들은 통제 목적으로 포함하였다. 종속변수는 국내 숙박여행 여부, 국내 당일여행 여부, 해외여행 여부와 국내 숙박여행 횟수, 국내 당일여행 횟수, 해외여행 횟수를 사용하여 총 6 개의 추정식을 만들었다. 세가지 여행 여부 변수는 모두 0 과 1 로 이루어진 더미변수로서 1 은 ‘여행을 가본적이 있다’ 를 의미한다. 종속변수로서 더미변수가 사용되었으므로 stata 에서 vce(robust) 명령어를 사용하여 이분산성 문제를 해결했다.

(여행 횟수 변수를 종속변수로 사용할 때에도 더욱 엄밀한 분석을 위하여 vce(robust) 명령어를 사용했다).

### 5.1.2 회귀 분석 결과

Independent Variables	Domestic Multi_days	Domestic Single_day	Overseas	Count Domestic Multi_days	Count Domestic Single_day	Count Overseas
Subjective Income Satisfaction(BL = 3, Moderate)						
1(Very Satisfied)	0.0277**	-0.0127	0.0374***	0.3371***	0.1158	0.1192***
2(Satisfied)	0.0286***	0.0050	0.0267***	0.1752***	0.1591***	0.0526***
4(Unsatisfied)	0.0014	0.0001	-0.0225***	-0.0391*	-0.0666	-0.0313***
5(Very Unsatisfied)	-0.0357***	-0.0397***	-0.0188***	-0.1073***	-0.2005***	-0.0253***
Working Time Satisfaction(BL= 3, Moderate)						
1(Very Satisfied)	0.0682***	0.0308***	0.0193***	0.2798***	0.5413***	0.0191
2(Satisfied)	0.0431***	0.0228***	0.0185***	0.1751***	0.1696***	0.0296***
4(Unsatisfied)	0.0105**	0.0008	-0.0081*	0.0060	-0.0117	-0.0154**
5(Very Unsatisfied)	-0.0018	-0.0251***	-0.0063	-0.0344	-0.0871	-0.0152
Constant						
	0.7535***	0.2631***	0.7463***	3.0479***	1.5605***	1.1345***
Observations	69,401	69,401	69,401	69,401	69,401	69,401
R-squared	0.1613	0.0870	0.1640	0.0942	0.0600	0.0855

Robust standard errors in brackets

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

\*k=1,000 BL= Baseline

우리의 주된 관심은 주관적 소득 만족도와 근로시간 만족도가 여행 횟수의 변화와 여행을 가는지 안가는지의 여부에 어떤 영향을 미치는지 확인하는 것이며, 그것에 대한 OLS 회귀 분석 결과는 위의 표와 같다. 결론적으로, <5.1.1>의 추정식에 나오듯이 다양한 변수들을 통제했을 때, 소득에 만족할수록 여행을 더 많이 가고, 만족하지 않을수록 여행을 덜 간다. 근로시간에 만족할수록 여행을 더 많이 가고, 만족하지 않을수록 여행을 덜 간다.

그것을 위의 표에서 확인할 수 있는데, 주관적 소득 만족도 설문에 ‘보통’ 이라고 응답한 사람에 비해 ‘매우 만족한다’, ‘약간 만족한다’ 라고 응답한 사람들이 국내 숙박 여행을 갈 확률이 각각 2.77%p, 2.86%p 올라가는 것으로 확인되었으며 이는 각각 5%, 1% 유의수준에서 유의했다. 또 소득에 ‘매우 불만족 한다’고 응답한 사람들은 국내 숙박 여행을 갈 확률이 3.57%p 낮아졌으며 이는 1% 유의수준에서 유의했다. 해외여행과 국내 당일 여행에 대해서도 비슷한 경향이 보인다.

근로시간 만족도도 여행 여부에 유의미한 영향을 미치는 것으로 확인되었다. 근로시간에 만족하냐는 설문에 ‘보통이다’라고 응답한 경우에 비해 ‘매우 만족한다’고 응답한 경우 국내 숙박 여행을 갈 확률이 6.82%p, 해외 여행을 가본적이 있다고 응답할 확률이 1.93% 증가했으며 둘다 1% 유의수준에서 유의했다. 마찬가지로 해외여행과 국내 당일 여행에 대해서도 비슷한 경향이 보인다.

여행 횟수에 대해서도 주관적 소득 만족도와 근로시간 만족도는 유의미한 영향을 미쳤다. 소득에 매우 만족하는 경우 보통이라 응답한 경우에 비해 국내숙박여행은 0.3371 회,

해외여행은 0.1191 회 더 많이 가는 것으로 나타났고 소득에 매우 불만족하는 경우 국내숙박 0.1073 회, 해외 0.0253 회로 여행을 덜 가는 것으로 나타났다.

근로시간에 매우 만족하는 경우 보통이라 응답한 경우에 비해 국내 숙박여행 0.28 회, 근로시간에 약간 만족하는 경우 보통이라 응답한 경우에 비해 국내숙박여행 0.175 회, 해외 여행 0.03 회 더 많이 가는 것으로 나타났으며 모두 1% 유의수준에서 유의했다.

Independent Variables	Domestic Multi_days	Domestic Single_day	Overseas	Count Domestic Multi_days	Count Domestic Single_day	Count Overseas
Year (BL = 2019)						
2013	-0.0337***	0.0376***	-0.1069***	-0.2406***	0.0026	-0.2108***
2015	-0.0202***	0.0109*	-0.0838***	-0.0921***	0.0705	-0.1703***
2017	0.0055	0.0239***	-0.0409***	0.0637*	0.2586***	-0.0740***
2021	-0.2768***	-0.2098***	-0.3598***	-0.9605***	-1.4719***	-0.5103***
2023	-0.0243***	-0.0618***	-0.2170***	0.0526	-0.4486***	-0.3211***
Sex(BL = Male)	0.0294***	0.0336***	0.0246***	0.0617***	-0.0348	0.0787***
Age(25~65, Continuous)	-0.0146***	-0.0029*	-0.0174***	-0.0620***	-0.0442***	-0.0319***
Education(BL = 3, High)						
0(Uneducated)	-0.1444***	-0.1392***	-0.1163***	-0.1853	-1.0362***	-0.1500***
1(Elementary)	-0.0978***	-0.0838***	-0.0391***	-0.1696***	-0.6280***	-0.0756***
2(Junior High)	-0.0682***	-0.0413***	-0.0219***	-0.1780***	-0.3934***	-0.0521***
4(Junior College)	0.0706***	0.0391***	0.0387***	0.3065***	0.3511***	0.0624***
5(Bachelor)	0.1066***	0.0520***	0.0874***	0.4281***	0.5110***	0.1210***
6(Master)	0.1292***	0.0793***	0.1605***	0.5658***	0.9580***	0.1841***
7(Doctoral)	0.1115***	0.0727***	0.2502***	0.5246***	0.8154***	0.2007***
Household Income (BL = 1, Under 100k)						
2(100k~200K)	0.0253**	0.0466***	-0.0005	0.0472	0.2017**	0.0042
3(200k~300k)	0.0770***	0.0871***	0.0250***	0.1574***	0.4638***	0.0540***
4(300k~400k)	0.1224***	0.1194***	0.0533***	0.3591***	0.8066***	0.0925***
5(400k~500k)	0.1677***	0.1447***	0.0922***	0.5345***	0.9062***	0.1384***
6(500k~600k)	0.2001***	0.1471***	0.1190***	0.7605***	1.0891***	0.1956***
7(Over 600K)	0.2341***	0.1685***	0.1543***	0.9728***	1.4515***	0.2744***
Household Size(BL =1)						
2	-0.0354***	-0.0072	-0.0052	-0.2332***	-0.2854***	0.0031
3	-0.0613***	-0.0149**	-0.0523***	-0.4319***	-0.4032***	-0.0853***
4	-0.0509***	-0.0124*	-0.0795***	-0.3456***	-0.3164***	-0.1274***
Marital Status(BL =1, Unmarried)						
2(Married)	0.1244***	0.0349***	0.0438***	0.4974***	0.7334***	0.0233**
3(bereaved)	0.0862***	0.0309**	0.0227**	0.4342***	0.7476***	-0.0106
4(Divorced)	0.0482***	-0.0099	0.0046	0.3311***	0.4879***	-0.0153
Disabled_welfarecard(BL = Yes)	0.0070	0.0057	0.0158**	-0.1096	-0.0621	0.0074
Employment Status(BL = 1, Regular Worker)						
2(Contract Worker)	-0.0347***	-0.0075	-0.0274***	-0.0925***	-0.0831	-0.0413***
Employment Stability(BL=1, Very Unstable)						
2(Unstable)	0.0051	0.0175***	-0.0104**	-0.0421	-0.0140	-0.0272***
3(Stable)	0.0104*	0.0183***	-0.0048	0.0015	-0.0068	-0.0181*
4(Very Stable)	0.0236***	0.0030	0.0018	0.1685***	0.1853**	-0.0113
Subjective Income Satisfaction(BL = 3, Moderate)						
1(Very Satisfied)	0.0277**	-0.0127	0.0374***	0.3371***	0.1158	0.1192***
2(Satisfied)	0.0286***	0.0050	0.0267***	0.1752***	0.1591***	0.0526***
4(Unsatisfied)	0.0014	0.0001	-0.0225***	-0.0391*	-0.0666	-0.0313***
5(Very Unsatisfied)	-0.0357***	-0.0397***	-0.0188***	-0.1073***	-0.2005***	-0.0253***
Working Time Satisfaction(BL= 3, Moderate)						
1(Very Satisfied)	0.0682***	0.0308***	0.0193***	0.2798***	0.5413***	0.0191
2(Satisfied)	0.0431***	0.0228***	0.0185***	0.1751***	0.1696***	0.0296***
4(Unsatisfied)	0.0105**	0.0008	-0.0081*	0.0060	-0.0117	-0.0154**
5(Very Unsatisfied)	-0.0018	-0.0251***	-0.0063	-0.0344	-0.0871	-0.0152
Constant	0.7535***	0.2631***	0.7463***	3.0479***	1.5605***	1.1345***
Observations	69,401	69,401	69,401	69,401	69,401	69,401
R-squared	0.1613	0.0870	0.1640	0.0942	0.0600	0.0855

Robust standard errors in brackets

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

\*k=1,000 BL= Baseline

추가로 F-test 를 진행 했을때, 종속 변수가 여행 여부(Binary Variable)일때는 근로시간 만족도가 더 큰 영향을 미치고 종속변수가 여행 횟수(Count Variable)일때는 소득 만족도가 더 큰 영향을 미친다. 다만 이것은 국내 숙박과 해외 여행의 경우이고, 국내 당일 여행은 앞서 말한 경향이 반대로 나타난다(여행 여부에 소득 만족도가 더 큰 영향, 여행 횟수에 근로시간 만족도가 더 큰 영향). 아래의 F-test 검정치를 비교해 봄으로서 위의 사실을 확인할 수 있다.

### 5.1.3 시사점

위와 같은 결과에 비추어 보았을 때 소득에 대한 만족도와 근로시간에 대한 만족도는 여행을 같 것인지 말 것인지, 또 간다면 몇번을 갈 것인지를 결정하는데 상당히 중요한 역할을 하고 있는 것으로 생각할 수 있다. 이러한 여행과 소득, 여행과 근로시간 간의 양의 상관관계는 숙박여행, 해외 여행에서 더욱 크고 뚜렷하게 나타난다. 이것을 확인함으로써, 관광업 진흥을 위한 정책을 결정할 때 노동자의 근로 환경 개선과 연관지어 정책 결정을 할 수 있는 실마리를 제공한 것이라고 생각된다. 그러나 위와 같은 OLS 분석은 본 분석에서 사용한 종속변수들의 특성(binary&count variable 이며 Zero-inflated 된 상태)에 기반했을때 적합한 추정방법이 아닌 것으로 판단되며 따라서 Logit 을 이용한 ZINB 모델로서 보다 정교한 분석을 수행하겠다.

## 5.2. ZINB 분석

### 5.2.1. 회귀식 설정

일반적인 ZINB 모델에서는 표본이 모두 0 인 경우인지를 결정짓는 Inflation 모델과 0 이 아닌 경우 음이항 분포를 따르는 Count 모델의 회귀식을 따로 결정해주는 것이 일반적이나 이 연구에서는 표본이 많고 통제 변수가 잘 되어있다는 점을 감안하여 두 회귀식이 동일하다는 것으로부터 회귀식을 설정하였다.

## ZINB MODEL

$$\mu = \beta_1 * \text{year} + \beta_2 * \text{sex} + \beta_3 * \text{age} + \beta_4 * \text{education} + \beta_5 * \text{household income} + \beta_6 * \text{household size} + \beta_7 * \text{marital status} + \beta_8 * \text{disabled} + \beta_9 * \text{satisfy\_wage\_worker} + \beta_{10} * \text{age\_square} + \beta_{11} * \text{Employ Stability} + \beta_{12} * \text{Income Satisfaction} + \beta_{13} * \text{Working Time Satisfaction}$$

$$\lambda = \beta_1 * \text{year} + \beta_2 * \text{sex} + \beta_3 * \text{age} + \beta_4 * \text{education} + \beta_5 * \text{household income} + \beta_6 * \text{household size} + \beta_7 * \text{marital status} + \beta_8 * \text{disabled} + \beta_9 * \text{satisfy\_wage\_worker} + \beta_{10} * \text{age\_square} + \beta_{11} * \text{Employ Stability} + \beta_{12} * \text{Income Satisfaction} + \beta_{13} * \text{Working Time Satisfaction}$$

### 5.2.2. 회귀 분석 결과

해외여행의 경우에는 먼저 해외여행을 가게되는 요인에는 금전적인 요인은 유의수준 10% 이내에서 유의미하지 않는 것으로 나타났다. 하지만 시간적인 요인으로는 유의수준 5%이내에서 시간적으로 여유로운 사람은 시간적으로 보통 여유로운 사람보다 여행을 한번이라도 가게 되는 것에 유의미한 결과로 나타났다. 하지만 여유롭지 않은 사람은 시간적 여유도가 보통인 사람보다는 유의수준 10% 이내에서 유의미하지 않은 것으로 나타났다. 이는 정도와는 상관없이 시간적으로 여유롭지 않은 사람들은 해외여행을 가지 않는 요인을 제공한다는 것을 이야기한다. 정량적으로 해석해 보았을 때에는 시간적 만족도가 매우 만족인 사람들은 보통인 사람들보다 여행을 갈 확률이 약 5.78% 더 높고 매우 만족하지 않은 사람은 보통인 사람들보다 여행을 갈 확률이 약 2.77% 더 낮음을 알 수 있다. 해외여행의 횟수에 대해서는 유의수준 1% 이내에서 금전적인 요인이 유의미적인 영향을 미친다. 특히 금전적인 수준이 올라가면 올라갈수록 여행 횟수에 증가량이 증가하는 효과를 시사한다. 소득 만족감이 매우 높은 사람은 보통인 사람보다 약 0.15 회 더 많이 가는 것으로 나타났다. 반면에 소득 만족감이 매우 낮은 사람은 보통인 사람보다 약 0.05 회 덜 가는 것으로 나타났다.

국내여행의 경우에도 비슷한 경향성을 확인하였다. 국내 숙박여행의 경우 해외여행을 가게되는 요인에는 마찬가지로 시간적인 요인으로 시간적 만족감이 매우 높은 사람은 보통인 사람보다 약 7.6% 더 높게 나왔으며 만족감이 매우 낮은 사람은 보통인 사람보다 약 0.6% 낮은 것으로 나타났다. 여행 횟수와 관련해서는 금전적인 요인으로 소득 만족감이 매우 높은 사람은 보통인 사람보다 약 0.3 회 많이 가는 것으로 나타났다. 국내여행 또한 여행 유무에 유의미한 영향을 미치는 데에는 시간적인 요인이 크게 작용했으며 여행 횟수에는 금전적인 요인이 크게 작용하였다. 이를 비교하기 위해서는 F-test 를 진행하였으며 다음 단락에서 설명할 예정이다. 다만 해외여행이랑 비교하였을 때, 이러한 요인들이 국내여행이 좀 더 횟수 또는 유무에 더욱 탄력적인 결과이다. 국내여행 내에서도 숙박여행 보다는 당일여행이 여행 횟수 및 유무에 더 탄력적인 결과를 이야기한다.

Probability of Not Going on International Trip

Independent Variable	dy/dx	Standard Error	Z-value	P-Value
Income Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.0376535	0.0327272	1.15	0.250
2(Satisfied)	-0.0067441	0.138679	-0.49	0.627
4(Unsatisfied)	0.0252875	0.0154979	1.63	0.103
5(Very Unsatisfied)	0.0248246	0.010431	1.47	0.140
Working Time Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	-0.0578645	0.0199398	-2.9	0.004
2(Satisfied)	-0.0286892	0.0144155	-1.99	0.047
4(Unsatisfied)	0.0125168	0.0163766	0.76	0.445
5(Very Unsatisfied)	0.0277792	0.0214868	1.29	0.196

Count of International Trips

Independent Variable	dy/dx	Standard Error	Z-value	P-Value
Income Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.1511536	0.0281048	5.38	0.000
2(Satisfied)	0.0772957	0.0092785	8.33	0.000
4(Unsatisfied)	-0.05500219	0.0069824	-7.16	0.000
5(Very Unsatisfied)	-0.0603852	0.010431	-5.79	0.000
Working Time Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.0312125	0.0103238	3.02	0.002
2(Satisfied)	0.0333658	0.007706	4.33	0.000
4(Unsatisfied)	-0.0168927	0.0082992	-2.04	0.042
5(Very Unsatisfied)	-0.0077153	0.118815	-0.65	0.516

Probability of Not Going on Domestic Single Day Trip

Independent Variable	dy/dx	Standard Error	Z-value	P-Value
Income Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.0119489	0.0155462	0.77	0.442
2(Satisfied)	-0.0096635	0.0070695	-1.37	0.172
4(Unsatisfied)	0.0013266	0.0062327	0.21	0.831
5(Very Unsatisfied)	0.0602736	0.009288	6.49	0.000
Working Time Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	-0.0241896	0.0090444	-2.67	0.007
2(Satisfied)	-0.0292384	0.0063692	-4.59	0.000
4(Unsatisfied)	-0.0036502	0.0075012	-0.49	0.627
5(Very Unsatisfied)	0.0306275	0.0105423	2.91	0.004

Count of Domestic Single Day Trips

Independent Variable	dy/dx	Standard Error	Z-value	P-Value
Income Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.3054045	0.1269468	2.41	0.016
2(Satisfied)	0.2533372	0.0546964	4.63	0.000
4(Unsatisfied)	-0.144118	0.0435639	-3.31	0.001
5(Very Unsatisfied)	-0.317695	0.0701145	-4.53	0.000
Working Time Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.5723991	0.0768647	7.45	0.000
2(Satisfied)	0.2209547	0.0469453	4.71	0.000
4(Unsatisfied)	-0.048051	0.0522023	-0.92	0.357
5(Very Unsatisfied)	-0.0754061	0.072484	-1.04	0.298

Probability of Not Going on Domestic Multi-Days Trip

Independent Variable	dy/dx	Standard Error	Z-value	P-Value
Income Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	-0.0181214	0.0174686	-1.04	0.300
2(Satisfied)	-0.0307146	0.0075254	-4.08	0.000
4(Unsatisfied)	-0.0025979	0.0069888	-0.37	0.710
5(Very Unsatisfied)	0.0476	0.0107118	4.44	0.000
Working Time Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	-0.0766253	0.0098378	-7.79	0.000
2(Satisfied)	-0.055018	0.0071892	-7.65	0.000
4(Unsatisfied)	-0.010908	0.0080648	-1.35	0.176
5(Very Unsatisfied)	-0.0068693	0.0121647	-0.56	0.572

Count of Domestic Multi-Days Trips

Independent Variable	dy/dx	Standard Error	Z-value	P-Value
Income Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.4541235	0.0798578	5.69	0.000
2(Satisfied)	0.2516589	0.028193	8.93	0.000
4(Unsatisfied)	-0.1169629	0.0242525	-4.82	0.000
5(Very Unsatisfied)	-0.2238867	0.0357091	-6.27	0.000
Working Time Satisfaction(Baseline: 3, Moderate)				
1(Very Satisfied)	0.3037509	0.0380863	7.98	0.000
2(Satisfied)	0.205123	0.0250801	8.18	0.000
4(Unsatisfied)	-0.0107194	0.0287569	-0.37	0.709
5(Very Unsatisfied)	-0.0326422	0.0395055	-0.83	0.409

F Test Results

해외여행

Likelihood of Engaging on International Trip	
Income Satisfaction	Working Time Satisfaction
Chi2( 04) = 10.57	Chi2( 04) = 38.08
Prob > Chi2 = 0.0319	Prob > Chi2 = 0.0000
Count of Engaging on International Trip	
Income Satisfaction = 0	Income Satisfaction = 0
Chi2( 04) = 161.43	Chi2( 04) = 18.72
Prob > Chi2 = 0.0000	Prob > Chi2 = 0.0022

국내 숙박여행

Likelihood of Engaging on Domestic Multi-Days Trip	
Income Satisfaction	Working Time Satisfaction
Chi2( 04) = 60.91	Chi2( 04) = 129.59
Prob > Chi2 = 0.0000	Prob > Chi2 = 0.0000

Count of Engaging on Domestic Multi-Days Trip	
Income Satisfaction	Working Time Satisfaction
Chi2( 04) = 193.28	Chi2( 04) = 58.40
Prob > Chi2 = 0.0000	Prob > Chi2 = 0.0000

국내 당일여행

Likelihood of Engaging on Domestic Single-Day Trip	
Income Satisfaction	Working Time Satisfaction
Chi2( 04) = 62.96	Chi2( 04) = 59.64
Prob > Chi2 = 0.0000	Prob > Chi2 = 0.0000

Count of Engaging on Domestic Single-Day Trip	
Income Satisfaction	Working Time Satisfaction
Chi2( 04) = 65.55	Chi2( 04) = 77.80
Prob > Chi2 = 0.0000	Prob > Chi2 = 0.0000

F-test 결과 해외여행, 국내 숙박여행과 해외 여행의 Extensive-Margin 에 대해서, 근로시간 만족도 변수에 대한 F 값이 주관적 소득 만족도에 대한 F 값보다 더 높기 때문에 국내 숙박여행과 해외 여행의 Extensive-Margin 을 결정하는것에는 근로시간에 대한 만족도가 더 중요한 요인이라고 할 수 있다. 국내 숙박여행과 해외 여행의 Intensive-Margin 에 대해서는, 주관적 소득 만족도에 대한 F 값이 근로시간 만족도에 대한 F 값보다 더 높다. 따라서 국내 숙박여행과 해외 여행의 Intensive-Margin 을 결정하는 것에는 주관적 소득 만족도가 더 주요한 요인이라고 해석할 수 있다. 다만 국내 당일여행에 대한 Extensive-Margin, Intensive-Margin 을 결정하는 메커니즘은 위에서 언급한 해외여행과 국내 숙박여행에 대한 그것과 정확히 반대로 나타나는 것이 관찰되었는데, 이것은 국내 당일여행에서는 또 다른, 새로운 패턴이 나타나고 있다는 것으로 유추해 볼 수 있다.

### 5.2.3. 시사점

회귀 결과에 미루어 보았을 때 여행 자체를 가게 되는 요인에는 시간적인 요인이 많은 영향을 미치지만 여행 횟수와 관련된 요인에는 금전적인 요인이 많은 영향을 미친다는 것을 알 수 있었다. 또한 여행 유무 요인에는 금전적인 요인이 거의 유의미하지 않은 반면에 여행 횟수 요인에는 시간적인 요인도 어느정도 영향을 미친다는 것을 알 수 있다. 현재 영과잉으로 나타나는 데이터 분포를 고려하면 여행 산업을 이끌어내기 위해서는 직장 내 금전적인 요인보다는 시간적인 요인을 고려하여 휴가 또는 기타 시간적 근무 여건을 향상시킨다면 여행 인구수가 크게 증가할 것을 기대한다.

## 6. 결론

본 연구는 대한민국 임금 근로자의 근로 시간에 대한 만족도의 변화와, 주관적 소득에 대한 만족도의 변화가, 국내외 여행(특히 국내 숙박여행과 해외여행)에 대한 Extensive-Margin 과 Intensive-Margin 에 영향을 미치는 정도의 차이를 한번에 비교하여 분석하기 위해서 ZINB 모델을 사용하였다. ZINB 모델 분석 결과, 국내 숙박여행과 해외 여행의 Extensive-Margin 에 대해서, 근로시간 만족도가 주관적 소득 만족도보다 더 큰 영향을 미치고, 국내 숙박여행과 해외 여행의 Extensive-Margin 을 결정하는것에는 근로시간에 대한 만족도가 더 큰 영향을 미친다는 결론을 얻을 수 있었다. 이것은 단순한 OLS 분석으로는 수행할 수 없는 비교분석으로, 한 모델안에서 한번에 두가지 만족도 변수에 대한 국내외 여행의 Extensive-Margin 과 Intensive Margin 을 비교하는 것을 가능하게 한다는 점에서 ZINB 모델을 사용한 의의가 있다.

이번 연구에서는 주관적 소득 만족감, 주관적 시간 만족감 등 임금 근로자에 관련된 데이터들을 활용하여 근로여건과 여행의 관련성을 알아보았다. 다만 다른 근로 유형인 자영업자, 일용직 근로자에 대해서는 어떤 요인이 여행 횟수 및 유무에 영향을 미치는지에서는 알 수 없다는 한계점이 있다. 이는 현재 수집한 데이터 및 데이터 전처리 이후 자영업자와 같은 근무유형과 관련된 근로 데이터가 없기 때문에 나타난 한계점이다. 만약, 임금근로자가

아닌 다른 근로 유형에 대한 근로 및 여행 데이터가 있다면 이를 활용하여 임금 근로자가 아닌 근로 여건과 여행과 관련된 연구를 진행할 수 있을 것이다.

## 7. 부록

### 7.1. 포아송 분포(Poisson Distribution)

포아송 분포는 단위 시간 안에 어떤 사건이 몇 번 발생할 것인지를 표현하는 이산 확률 분포이다.

정해진 시간 안에 어떤 사건이 일어날 횟수에 대한 기대값을  $\lambda$ 라고 했을 때, 그 사건이  $k$ 회 일어날 확률은 다음과 같다.

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

여기서  $e$ 는 자연상수이다.

포아송 분포는 다음과 같은 특성을 갖는다.

어떤 단위구간(예, 1 일)동안 이를 더 짧은 작은 단위의 구간(예: 1 시간)로 나눌 수 있고 이러한 더 짧은 단위구간 중에 어떤 사건이 발생할 확률은 전체 척도 중에서 항상 일정해야 한다.

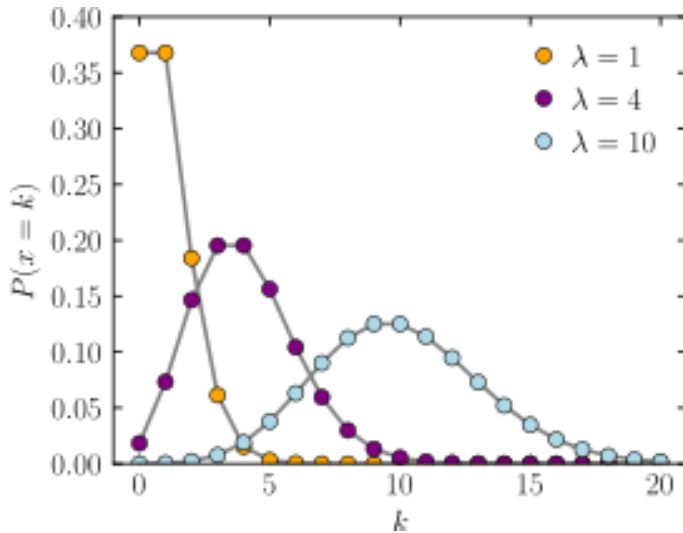
두 개 이상의 사건이 동시에 발생할 확률은 0에 가깝다.

어떤 단위구간의 사건의 발생은 다른 단위구간의 발생으로부터 독립적이다.

특정 구간에서의 사건 발생확률은 그 구간의 크기에 비례한다.

포아송분포 확률 변수의 기대값과 분산은 모두  $\lambda$ 이다.

이항분포를 따르는 위와 같은 확률변수  $X$ 에서,  $n$ 이 대단히 크고  $p$ 가 대단히 작을 경우, 이 확률변수  $X$ 는  $\lambda=np$ 인 포아송 분포로 근사할 수 있다.



## 7.2. 음이항 분포(Negative Binomial Distribution)

음이항 분포는 이항 분포에서부터 기인한다 이항 분포가 성공 확률이  $p$  인 사건을  $n$  번 실행하였을 때, 성공 횟수의 분포라고 한다면, 음이항 분포는  $k$  번의 성공을 얻기 위하여 실행한 실험 횟수의 분포라고 할 수 있다.

이를 식으로 표현하면 다음과 같다.

$$b(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}, x = k, k+1, k+2, \dots$$

### 참고자료

통계청. MDIS. 사회조사(복지/사회참여/문화와여가/소득과소비/노동)데이터.

통계청.KOSIS 국가통계포털(문화여가/국민여행조사/국내여행)데이터

통계청. 지표누리(국가발전지표/고용과 노동/일자리만족도/근로시간)데이터

OECD.Stat >Labour Force Statistics>Weekly Worked Hours

(Kto) 2024 관광트렌드 전망 및 분석 보고서 , 한국관광공사, 23.12, 23-26

이강욱 김희수(2001). 주 5 일제 근무제 도입에 따른 관광정책 대응방안. 한국관광연구원

Woo Gon Kim(2002). 국내 여행 지출에 영향을 미치는 요인 분석 연구.

문화체육관광부. 근로자 휴가조사(2023)

<https://m.blog.naver.com/goldmonji2/220726434215>