

## 한국의 TV 시청유형 분석 및 시청자 유형별 시청률 예측 모형\*

정진욱\*\*

---

---

TV 프로그램의 성격과 시청자의 사회·경제적 특성에 따른 채널 선택 유형을 분석하고 시청률의 결정요인을 실증적으로 연구하는 것은 광고매체의 효율적 선택과 광고 효과의 정확한 예측을 위하여 매우 중요한 작업이다. 이 연구는 회귀분석을 통하여 우리나라 TV 시청자들의 프로그램 선택에 영향을 끼치는 요인들을 식별하고 그 영향의 정도를 추정함으로써 38개 시청자 유형별 시청률 예측모형을 구축한다. 분석 결과 프로그램의 시청점유율에 영향을 미치는 요인은 프로그램 유형, 방영시간대, 프로그램의 질, 채널충성도 등인 것으로 식별되었으나, 시청률 결정요인 및 모형의 예측력은 시청자 유형별로 매우 상이하게 나타났다.

---

---

### I. 연구의 목적과 필요성

상업광고를 방영하는 TV 채널이 늘어나고 케이블 TV와 위성방송이 도입됨

---

\* 본 연구를 위하여 방송학과 관련한 자문을 해 주신 조정식 교수(중앙대학교 신문방송학과)와 시청률 자료를 제공해 주신 MSK(Media Service Korea)와 동방커뮤니케이션스에 깊은 감사를 드린다.

\*\* 연세대학교 상경대학 경제학과, 서울특별시 서대문구 신촌동 134, 120-749.

에 따라, 시청자들의 TV 시청 유형과 시청률 결정요인에 대한 방송사와 광고주의 관심은 날로 증가하고 있다. 프로그램의 성격과 시청자의 사회·경제적 특성에 따른 채널 선택 유형을 분석하고 시청률의 결정요인을 실증적으로 연구하는 것은 광고매체의 효율적 선택과 광고 효과의 정확한 예측을 위하여 매우 중요한 작업이다. 그러나 우리나라의 경우 시청률의 측정은 비교적 폭넓게 시도되고 있는 반면, 시청자의 채널 선택 과정이나 시청률 결정요인에 대한 실증적 분석 및 예측은 거의 없었다.

이 연구는 첫째, 우리나라 TV 시청자들의 프로그램 선택에 영향을 끼치는 요인들을 식별하고 그 영향의 정도를 추정하며, 둘째, 시청자 유형별로 시청률 예측모형을 구축함으로써 새로운 TV 프로그램이 방영될 때 그 시청률을 시청자 유형별로 예측할 수 있게 하는 것을 목적으로 한다.

## II. 시청률 결정요인 분석에 관한 기존의 연구

시청자의 TV 시청행위를 설명하는 모형은 크게 두 가지로 구분할 수 있다. 하나는 개별 시청자의 채널 선택이 어떤 요인들에 의해 결정되는가를 분석하는 개인모형(individual model)이고, 다른 하나는 개별 프로그램 또는 채널의 시청률이 어떤 요인들에 의해 결정되는가를 분석하는 통합모형(aggregate model) 또는 시청률모형(ratings model)이다.

### 1. 개인모형

개인모형은 효용극대화(utility maximization) 원칙에 의거한 이산형 선택모형(discrete choice model)으로 설정된다.  $m$ 개의 채널이 선택 가능할 때, 이산형 선택모형은 다음과 같은 식으로 표현될 수 있다.

$$C_i = X_i\beta + u_i$$

여기서  $C_i$ 는  $i$  시청자가 선택한 채널(0, 1, 2, ...,  $m$  중의 한 숫자로 표시됨),  $X_i$ 는  $i$  시청자의 사회·경제적 특성을 나타내는 변수들 및 해당 채널의 프로그램 특성을 나타내는 변수들로 이루어진, 시청자의 채널선택 행위를 설명할 수 있는 설명변수 벡터(예를 들어, 나이, 교육수준, 직업, 소득, 프로그램의 종류, 채널의 특성 등이 이에 포함됨),  $u_i$ 는 관측불가능한 요인들을 나타내는 오차항(errors)이다.

개인모형의 추정은 최우추정법(maximum likelihood estimation)으로 이루어진다. 우도함수(likelihood function)의 도출은 채널선택 행태를 어떻게 설정하는가에 따라 달라질 수 있다. Tavakoli and Cave [21], Rust and Alpert [18], Rust and Eechambadi [19] 등의 기존 연구들은 대개 이 '시청결정'과 '채널선택'의 두 단계 결정이 동시에 일어난다고 단순화하고 '다원 로짓 최우추정법'(multinomial logit MLE)으로 개인모형을 추정한다. 즉, 시청자가 0(TV 시청 없음), 1(채널 A), 2(채널 B), ...,  $m$ (채널 M) 중 하나를 단번에 선택한다는 것이다. 의사결정 과정을 이렇게 단순화하면 개인모형은 다원 로짓 최우추정법으로 비교적 쉽게 추정될 수 있다.<sup>1)</sup>

만일 시청자의 의사결정 과정이 우선 TV를 시청할 것인가 말 것인가를 결정하는 단계와 시청한다면 어떤 채널/프로그램을 시청할 것인가를 결정하는 단계의 두 단계로 이루어진다면 개인모형은 '연쇄결정모형'(sequential response model)으로 추정될 수 있으며 이 경우의 우도함수는 다원 로짓의 경우와는 다르게 설정되어야 한다.<sup>2)</sup>

1) 오차항의 분포를 정규분포로 가정하면 '다원 프로빗 최우추정법'(multinomial probit MLE)으로 추정하여야 하나, 추정을 용이하게 하기 위하여 오차항이 정규분포와 흡사한 '로지스틱 분포'(logistic distribution)라고 가정하고 다원 로짓 MLE를 사용하는 것이 일반적이다.

2) Maddala [14]를 참조하시오.

## 2. 시청률모형

개인모형은 개별 시청자가 어떤 의사결정 과정을 거쳐 채널/프로그램을 선택하느냐를 분석하는 반면, 시청률모형은 그 통합적 현상으로 나타나는 채널/프로그램의 시청률이 어떤 요인들에 의해 결정되는가를 분석한다. 시청률모형은 다음 회귀분석모형(regression model)으로 설정된다.

$$R_i = Z_i\alpha + e_i$$

여기서  $R_i$ 는  $i$ 채널/프로그램의 시청률,  $Z_i$ 는  $i$ 채널/프로그램의 특성을 나타내는 변수들로 이루어진, 시청률 설명변수의 벡터(예를 들어, 프로그램 성격, 방영시각, 지역, 나이, 채널의 특성 등이 이에 포함됨),  $e_i$ 는 관측 불가능한 요인들을 나타내는 오차항(errors)이다. 시청률모형은 최소자승법(least squares estimation)으로 추정된다.

Darmon [10], Henry and Rinne [13] 등은 대표적인 시청률모형 추정 연구이며, Gensch and Shaman [11]과 Gensch and Shaman [12]은 전통적인 회귀분석 대신 시계열모형(time series model)을 이용하여 시청행태를 분석하였다. 우리나라 시청자의 시청행태를 연구한 결과로는, 한혜경 [8], 심미선 [4], 박원기·김수영 [3], 그리고 안민호 [5] 등이 있다.

## Ⅲ. 모 형

이 연구는 시청률의 예측에 중점을 두고 있으므로 개인모형은 추정하지 않고, 시청률모형을 회귀분석으로 추정한다. 이 연구가 분석하는 모형은 다음과 같다.

$$S_i = \alpha + \sum_j \beta_j TD_{ji} + \sum_k \gamma_k CD_{ki} + \sum_m \delta_m SD_{mi} \\ + \theta_H GDH_i + \theta_L GDL_i + \phi DUR_i + u_i$$

모형에 사용된 변수들은 다음과 같다.

### 1. 점유율 ( $S_i$ )

종속변수로는 프로그램  $i$ 의 평균 시청점유율(share)  $S_i$ 를 사용한다. 시청점유율이란 TV를 본 시청자들 사이에서의 점유율을 말한다. 시청률의 분석 및 예측에 있어서 두 가지 방안을 고려할 수 있다. 하나는 실제 시청률을 종속변수로 하여 분석하는 것이고, 다른 하나는 점유율을 종속변수로 하여 분석한 후 시청가계비율(Household Using TV: HUT) 혹은 시청자비율(People Using TV: PUT)을 곱하여 실제 시청률을 산출하는 방법이다. 이 연구에서 점유율을 활용하는 이유는 점유율의 변화가 시청률에 비하여 상대적으로 안정적이며, HUT 및 PUT는 계절 등 상황에 따라 충분히 합리적으로 산출될 수 있다는 점을 고려한 것이다.

### 2. 프로그램 유형 가변수 ( $TD_i$ )

TV 프로그램의 유형이 점유율에 미치는 영향을 고려하기 위하여 설명변수로 프로그램 유형 가변수(Type Dummy)를 사용한다. 프로그램의 유형은 방송공사의 구분에 따라 다음 9개로 분류된다.

- TD1: 보도(05)
- TD2: 정보(10)
- TD3: 드라마·영화(15)
- TD4: 오락(20)
- TD5: 스포츠(25)

TD6:	교육(30)
TD7:	어린이·유아(35)
TD8:	종합(88)
TD9:	기타(99)

가변수함정(dummy trap)을 피하기 위하여 TD9을 제외하고 나머지 8개의 가변수를 설명변수로 사용하였다.

### 3. 채널 가변수 ( $CD_i$ )

채널 충성도 효과를 포착하기 위하여 채널 가변수(Channel Dummy)를 설명변수로 사용한다. 채널 가변수의 설정은 다음과 같다.

CD1:	KBS1(1)
CD2:	MBC(3)
CD3:	SBS(4)

### 4. 방영시간대 가변수 ( $SD_i$ )

방영시간대가 점유율에 미치는 영향을 포착하기 위하여 방영시간대 가변수(Slot Dummy)를 설명변수로 사용한다. 방영시간대의 구분은 다음과 같다.

WDDC:	평일 06:00~07:00
Wddb:	평일 07:00~08:30
WDMA:	평일 08:30~09:30
WDMB:	평일 09:30~13:00
WDAC:	평일 13:00~17:00
WDEB:	평일 17:00~19:00
WDEA:	평일 19:00~20:00
WDES:	평일 20:00~23:00

WDNA:	평일 23:00~24:00
WDNB:	평일 24:00~24:30
WDNC:	평일 24:30~
SATDC:	토요일 06:00~07:00
SATDB:	토요일 07:00~08:30
SATMA:	토요일 08:30~09:30
SATMB:	토요일 09:30~13:00
SATAB:	토요일 13:00~17:00
SATEA:	토요일 17:00~19:00
SATES1:	토요일 19:00~21:00
SATES2:	토요일 21:00~23:00
SATNA:	토요일 23:00~24:00
SATNB:	토요일 24:00~24:30
SATNC:	토요일 24:30~
SUNDC:	일요일 06:00~07:00
SUNDB:	일요일 07:00~08:30
SUNMA1:	일요일 08:30~11:00
SUNMA2:	일요일 11:00~13:00
SUNAA:	일요일 13:00~17:00
SUNEA:	일요일 17:00~19:00
SUNES1:	일요일 19:00~21:00
SUNES2:	일요일 21:00~23:00
SUNNS:	일요일 23:00~23:30
SUNNA:	일요일 23:30~24:00
SUNNB:	일요일 24:00~24:30
SUNNC:	일요일 24:30~

WDDC를 제외한 33개의 가변수가 설명변수로 사용되었다.

## 5. 등급 가변수 ( $GD_i$ )

TV 프로그램의 점유율에 가장 영향을 미치는 것은 프로그램의 질(quality)이

라고 할 수 있다. 이 연구에서는 프로그램의 질을 상·중·하 세 등급으로 나누어 분석한다. 등급분류는 여러 가지 방법으로 가능할 것이나, 이 연구에서는 전월의 평균점유율을 이용하여 전월의 평균점유율이 상위 25%에 속하면 '상' 등급, 하위 25%에 속하면 '하' 등급, 그리고 중앙 50%에 속하면 '중' 등급으로 규정하였다. 등급가변수(Grade Dummy)의 설정은 다음과 같다.

- 만일 '상' 등급이면  $GDH_i = 1$ , '상' 등급이 아니면  $GDH_i = 0$
- 만일 '하' 등급이면  $GDL_i = 1$ , '하' 등급이 아니면  $GDL_i = 0$

## 6. 방영시간 길이 ( $DUR_i$ )

프로그램의 길이가 점유율에 미치는 영향을 포착하기 위하여 프로그램 방영시간의 길이(duration)를 설명변수로 사용한다.

## IV. 자 료

이 연구는 우리나라의 대표적 시청률 조사기관인 MSK(Media Services Korea)의 피플미터(People Meter) 자료를 활용하였다. MSK의 시청률 조사는 서울지역 300가구를 대상으로 하여 자동으로 이루어지므로 이 연구도 서울지역 시청자로 그 영역이 제한된다. 이 연구가 사용한 자료는 1999년 11월 한 달간 지상파(KBS1, KBS2, MBC, SBS)로 방영된 총 3,033개 프로그램의 점유율 및 각 설명변수들이다.

1999년 11월에 방영된 3,033개 프로그램으로부터 산출된 점유율 및 각 설명변수들의 특성치는 <표 1>과 같다.



<표 1> 각 변수의 특성치

변 수	단 위	관찰치 수	평 균	표준편차	최소값	최대값
SHARE	%	3033	23.676560	11.098800	0	90
TD1	binary	3033	0.316189	0.465064	0	1
TD2	binary	3033	0.237389	0.425553	0	1
TD3	binary	3033	0.153314	0.360349	0	1
TD4	binary	3033	0.124299	0.329977	0	1
TD5	binary	3033	0.007913	0.088617	0	1
TD7	binary	3033	0.104517	0.305981	0	1
TD8	binary	3033	0.053742	0.225545	0	1
CHD1	binary	3033	0.237059	0.425349	0	1
CHD2	binary	3033	0.279591	0.448872	0	1
CHD3	binary	3033	0.251566	0.433985	0	1
WDDDB	binary	3033	0.056050	0.230056	0	1
WDMA	binary	3033	0.052753	0.223577	0	1
WDMB	binary	3033	0.105506	0.307255	0	1
WDAC	binary	3033	0.064293	0.245314	0	1
WDEB	binary	3033	0.123640	0.329225	0	1
WDEA	binary	3033	0.061325	0.239966	0	1
WDES	binary	3033	0.133202	0.339848	0	1
WDNA	binary	3033	0.035279	0.184514	0	1
WDNB	binary	3033	0.011210	0.105300	0	1
WDNC	binary	3033	0.036268	0.186986	0	1
SATDC	binary	3033	0.014837	0.120919	0	1
SATDB	binary	3033	0.008572	0.092205	0	1
SATMA	binary	3033	0.008902	0.093945	0	1
SATMB	binary	3033	0.027366	0.163173	0	1
SATAB	binary	3033	0.016815	0.128599	0	1
SATEA	binary	3033	0.013518	0.115497	0	1
SATES1	binary	3033	0.011869	0.108316	0	1
SATES2	binary	3033	0.013188	0.114099	0	1
SATNA	binary	3033	0.003297	0.057335	0	1
SATNC	binary	3033	0.004616	0.067795	0	1
SUNDC	binary	3033	0.014507	0.119588	0	1
SUNDB	binary	3033	0.011210	0.105300	0	1
SUNMA1	binary	3033	0.016485	0.127354	0	1
SUNMA2	binary	3033	0.012529	0.111247	0	1
SUNAA	binary	3033	0.018464	0.134643	0	1
SUNEA	binary	3033	0.011210	0.105300	0	1
SUNES1	binary	3033	0.013188	0.114099	0	1
SUNES2	binary	3033	0.012529	0.111247	0	1
SUNNS	binary	3033	0.000330	0.018158	0	1
SUNNA	binary	3033	0.002967	0.054402	0	1
SUNNC	binary	3033	0.004286	0.065339	0	1
GDH	binary	3033	0.243323	0.429160	0	1
GDL	binary	3033	0.203759	0.402858	0	1
GUR	분	3033	34.370590	22.102050	1	179

## V. 추정결과

이 연구는 새로운 TV 프로그램이 방영될 때 그 점유율을 예측하는데 중점을 두고 있다. 점유율은 시청자의 유형에 따라 달라질 것이다. 예를 들어, 새로운 시사프로그램이 방영될 때 그 점유율로 나타나는 시청자의 반응은 남·녀간, 연령대간, 직업간 차이가 있을 것이다. 이러한 시청자 유형별 차이를 점유율 예측에 반영하기 위해서는 점유율 결정모형을 여러 유형의 시청자 群에 대하여 추정해야 한다. 이 연구는 시청자를 다음의 38개 유형으로 구분하여 각각의 유형에 대하여 점유율 결정모형을 회귀분석하였다.

- 유형 1: 전체 시청자
- 유형 2: 남녀 구분 - 남자 전체
- 유형 3: 남녀 구분 - 여자 전체
- 유형 4: 학력별 구분 - 고졸 미만
- 유형 5: 학력별 구분 - 고졸 이상
- 유형 6: 수입별 구분 - 고소득
- 유형 7: 수입별 구분 - 중소득
- 유형 8: 수입별 구분 - 저소득
- 유형 9: 남자 연령별 구분 - 십대 미만(0~9)
- 유형 10: 남자 연령별 구분 - 십대(10~19)
- 유형 11: 남자 연령별 구분 - 이십대(20~29)
- 유형 12: 남자 연령별 구분 - 삼십대(30~39)
- 유형 13: 남자 연령별 구분 - 사십대(40~49)
- 유형 14: 남자 연령별 구분 - 오십대(50~59)
- 유형 15: 남자 연령별 구분 - 육십대 이상(~60)
- 유형 16: 여자 연령별 구분 - 십대 미만(0~9)
- 유형 17: 여자 연령별 구분 - 십대(10~19)
- 유형 18: 여자 연령별 구분 - 이십대(20~29)
- 유형 19: 여자 연령별 구분 - 삼십대(30~39)
- 유형 20: 여자 연령별 구분 - 사십대(40~49)

〈표 2〉 점유율 결정모형(전체 시청자)

Variables	Estimated Coefficient	Standard Error	t-value	P-value
INTERCEPT	13.8886	2.930110	4.74	0.0001
TD1	3.69853	2.817688	1.313	0.1894
TD2	1.341431	2.819382	0.476	0.6343
TD3	5.917396	2.846847	2.079	0.0377
TD4	2.672995	2.821154	0.947	0.3435
TD5	7.730948	3.236891	2.388	0.0170
TD7	4.066841	2.874887	1.415	0.1573
TD8	4.89493	2.943543	1.663	0.0964
CHD1	1.501792	0.468032	3.209	0.0013
CHD2	3.803163	0.447676	8.495	0.0001
CHD3	1.409163	0.438688	3.212	0.0013
WDDDB	0.293741	0.838803	0.35	0.7262
WDMA	1.715668	0.868589	1.975	0.0483
WDMB	2.875735	0.711924	4.039	0.0001
WDAC	2.212176	0.793389	2.788	0.0053
WDEB	2.144242	0.726828	2.95	0.0032
WDEA	1.743233	0.802840	2.171	0.0300
WDES	2.115702	0.693151	3.052	0.0023
WDNA	0.808617	0.943362	0.857	0.3914
WDNB	3.310474	1.432250	2.311	0.0209
WDNC	5.521606	0.927337	5.954	0.0001
SATDC	-0.10673	1.245213	-0.086	0.9317
SATDB	1.189402	1.615047	0.736	0.4615
SATMA	1.133809	1.599995	0.709	0.4786
SATMB	1.757844	1.014247	1.733	0.0832
SATAB	0.530648	1.227225	0.432	0.6655
SATEA	1.971019	1.403551	1.404	0.1603
SATES1	-0.20123	1.405947	-0.143	0.8862
SATES2	-1.98673	1.340088	-1.483	0.1383
SATNA	-3.74332	2.520642	-1.485	0.1376
SATNC	-3.51534	2.187643	-1.607	0.1082
SUNDC	4.227172	1.281623	3.298	0.0010
SUNDB	2.405986	1.419885	1.694	0.0903
SUNMA1	-0.83338	1.237700	-0.673	0.5008
SUNMA2	-1.16642	1.379705	-0.845	0.3979
SUNAA	-1.05435	1.184124	-0.89	0.3733
SUNEA	3.252601	1.451093	2.241	0.0251
SUNES1	0.034159	1.346639	0.025	0.9798
SUNES2	-0.97657	1.379009	-0.708	0.4789
SUNNS	-16.2635	7.760824	-2.096	0.0362
SUNNA	-2.23969	2.630164	-0.852	0.3945
SUNNC	-0.30191	2.217822	-0.136	0.8917
GDH	12.21214	0.381743	31.99	0.0001
GDL	-8.35212	0.437428	-19.094	0.0001
GUR	0.049414	0.007471	6.614	0.0001

주:  $R^2 = 0.5293$ ,  $n = 3,033$ .

유형 21:	여자 연령별 구분 - 오십대(50~59)
유형 22:	여자 연령별 구분 - 육십대 이상(~60)
유형 23:	남자 연령별 구분 - 기타(35~44)
유형 24:	남자 연령별 구분 - 기타(30~44)
유형 25:	남자 연령별 구분 - 기타(25~39)
유형 26:	남자 연령별 구분 - 기타(25~34)
유형 27:	남자 연령별 구분 - 기타(20~34)
유형 28:	남자 연령별 구분 - 기타(15~29)
유형 29:	남자 연령별 구분 - 기타(15~24)
유형 30:	여자 연령별 구분 - 기타(35~44)
유형 31:	여자 연령별 구분 - 기타(30~44)
유형 32:	여자 연령별 구분 - 기타(25~39)
유형 33:	여자 연령별 구분 - 기타(25~34)
유형 34:	여자 연령별 구분 - 기타(20~34)
유형 35:	여자 연령별 구분 - 기타(15~29)
유형 36:	여자 연령별 구분 - 기타(15~24)
유형 37:	주부 전체
유형 38:	주부 20~30대

〈표 2〉는 전체 시청자를 대상으로 점유율 결정모형을 회귀분석한 결과이다.

〈표 2〉를 통하여 우리는 프로그램의 시청점유율(전체 시청자)이 어떤 요인에 의하여 영향을 받는지 알 수 있다. 우선 프로그램 유형이 점유율에 미치는 영향을 살펴보면, 스포츠(TD5) 프로그램이 기타(TD9)에 비하여 평균적으로 7.73% 높은 점유율을 보이고 있으며, 드라마 및 영화(TD3)가 기타보다 5.92%, 종합(TD8) 프로그램이 기타보다 4.89% 높아 인기 있는 유형으로 나타난다. 회귀계수들의 통계적 유의도도 높다. 채널에 대한 선호도는 채널가변수(CD)로 포착되는데, KBS1이 KBS2에 비하여 평균 1.5% 높은 점유율을, MBC가 KBS2에 비하여 평균 3.8% 높은 점유율을, 그리고 SBS가 KBS2보다 평균 1.41% 높은 점유율을 보이고 있으며, 회귀계수들은 통계적으로 매우 유의하다.

방영시간대에 따른 점유율의 변화는 다른 변수들에 비하여 그 통계적 유의도가 상대적으로 낮은 편이다. 이는 첫째, 예측정확도를 높이기 위하여 방영시간대의 구분을 지나치게 세분화함으로써 변별력이 약해진 점과, 둘째, 방영시간대

가변수와 등급가변수(Grade Dummies) 사이에 존재하는 유의한 상관관계 때문인 것으로 판단된다. 그러나 이 연구의 주 목적은 미래 점유율의 예측에 있기 때문에 세분화된 방영시간대의 구분을 유지하기로 한다.

프로그램의 질(quality)을 나타내는 등급가변수(GDH, GDL)는 점유율이 프로그램의 질에 많은 영향을 받는다는 사실을 실증하고 있다. 등급이 '상'인 프로그램은 등급이 '중'인 프로그램에 비하여 평균적으로 12.2% 높은 점유율을 가지며, 등급이 '하'인 프로그램은 등급이 '중'인 프로그램에 비하여 평균적으로 8.35% 낮은 점유율을 갖는다는 것이 <표 2>에 나타난다. 프로그램의 길이(duration)도 점유율에 유의한 영향을 끼치는 것으로 나타나는데, 그 영향은 방영시간이 1분 늘어남에 따라 점유율이 0.05% 높아지는 것으로 실증되었다.<sup>3)</sup>

이 연구는 전체 시청자 외에, 위에 제시된 38개의 시청자 유형별로 점유율 결정모형을 회귀분석하였다. 추정결과는 시청자 유형별로 매우 상이하게 나타났으며, 각 모형의 예측력도 시청자 유형에 따라 상당히 달랐다. 각 유형별 추정결과는 지면을 절약하기 위하여 생략하나, 각 모형의 예측력은 제 VI 장에서 설명하는 의태분석에 포함되어 있음을 밝힌다.<sup>4)</sup>

## VI. 예 측

새로 방영되는 프로그램의 점유율은 점유율 결정모형에서 추정된 회귀계수(regression coefficients)들을 이용하여 예측된다. 예측치는 다음 식에 의하여 계산된다.

3) <표 2>에 제시된 회귀분석의 안정성을 점검하기 위하여 이분산성(heteroskedasticity), 다중공선성(multicollinearity) 등이 검정되었다. 검정결과 <표 2>의 결과는 이분산성을 갖고 있는 것으로 나타나, 추정가능 일반화 최소자승법(Feasible Generalized Least Squares : FGLS)으로 교정을 시도하였다. FGLS 추정결과는  $R^2$ 나 잔차(residuals)의 안정성면에서 우월하였으나, FGLS에 근거한 예측은 <표 2>에 제시된 최소자승추정법(OLS)보다 부정확하였다.

4) 요청이 있을 때, 저자는 언제든지 각 유형별 추정결과를 제공할 것이다.

$$S^* = \hat{\alpha} + \sum_j \hat{\beta}_j TD_j^* + \sum_k \hat{\gamma}_k CD_k^* + \sum_m \hat{\delta}_m SD_m^* \\ + \hat{\theta}_H GDH^* + \hat{\theta}_L GDL^* + \hat{\phi} DUR^*$$

여기서  $S^*$ 는 새 프로그램의 점유율 예측치를 나타내고,  $TD^*$ ,  $CD^*$ ,  $SD^*$ ,  $GDH^*$ ,  $GDL^*$ ,  $DUR^*$  등은 새 프로그램의 설명변수 값들을 나타낸다. 이 설명변수 값들 중 대부분은 방영시점과 프로그램의 내용을 알면 구할 수 있으나, 새 프로그램의 등급을 나타내는  $GDH$ 와  $GDL$ 은 주관적으로 판단·결정되어야만 한다.

예측치는 새 프로그램의 점유율을 가장 잘 추측하는 값이긴 하지만 실제 점유율이 예측치와 완전히 일치할 가능성은 수학적으로 0이다. 따라서 예측은 대개 예측치의 주위에 신뢰구간(confidence interval)을 설정하는 방식으로 이루어지는 것이 상례이다. 예측치  $S^*$ 의 예측구간은 신뢰도  $(1-\alpha)$ 에서 다음과 같이 산출된다.

$$\left[ S^* - t^*_{\frac{\alpha}{2}} \cdot \sqrt{\hat{\sigma}^2(1+x^{*'}(X'X)x^*)}, S^* + t^*_{\frac{\alpha}{2}} \cdot \sqrt{\hat{\sigma}^2(1+x^{*'}(X'X)x^*)} \right]$$

여기서,  $X$ 는 추정에 사용된 모든 설명변수들로 이루어진  $(n \times k)$ 행렬이며  $x^*$ 는 새로운 프로그램에 관한 설명변수들로만 이루어진  $(1 \times k)$ 행벡터(row vector)이다.

이 연구에서 추정된 모형에 근거한 예측이 얼마나 정확한지를 검증하기 위하여 의태분석(simulation)을 시행하였다. 즉, 1999년 11월 자료를 이용하여 1999년 12월 첫째 주 및 둘째 주의 프로그램(684개)의 점유율을 예측한 후, 1999년 12월의 실제 점유율이 그 예측치와 얼마나 차이가 나는지를 검증하였다.

의태분석 결과 예측은 비교적 정확한 것으로 나타났다. 90% 예측구간을 사용해서 예측한 경우, 실제 점유율의 93.57%(684개의 프로그램 중 640개)를 올바르게 예측하였으며, 95% 예측구간을 사용해서 예측한 경우에는 실제 점유율의 95.61%(684개의 프로그램 중 654개)를 올바르게 예측하였다. 예측오차(실제치와 예측치의 차이를 절대값으로 표시한 것)의 평균은 5.51%, 예측오차의 표준편차

는 5.09%로 나타났다. 그러나, 예측 구간의 폭은 큰 편이어서 90% 구간의 경우 평균  $\pm 12.7\%$ , 95% 구간의 경우에는  $\pm 15.1\%$ 에 달하였다. 이는 추정된 회귀모형의 표준오차(standard error of regression)가 높는데 기인하는 것으로 보이는데, 이를 개선하기 위해서는 제작비등 프로그램의 질(quality)을 반영할 수 있는 자료의 보완이 필요할 것으로 판단된다.

### ▣ 참고 문헌 ▣

1. 강대인, 「한국 텔레비전 편성의 특성과 전략에 관한 연구」, 고려대학교 신문방송학과 박사학위 논문, 1993.
2. 김동훈·윤기중·이봉운·조성호, 「시청률 조사 검증 연구」, 한국방송개발원, 1994.
3. 박원기·김수영, “시청률 예측에 관한 연구”, 한국방송공사, 2000.
4. 심미선, 「시청습관의 유형과 시청행위 연구」, 고려대학교 신문방송학과 박사학위 논문, 1999.
5. 안민호, “수용자의 TV 프로그램 시청 흐름과 프로그램간 연결구조에 관한 연구”, 숙명여자대학교, 2000.
6. 원우현·박기성·최현철·한균태·김용호, 「시청행태 조사의 검증기구 수립방안에 관한 연구」, 방송위원회, 1994.
7. 조성호, 「텔레비전 시청률 조사 연구」, 서울: 커뮤니케이션북스, 1998.
8. 한혜경, 「시청자 특성에 따른 시청패턴 분석」, 연세대학교 신문방송학과 박사학위 논문, 1998.
9. Amemiya, T., “Qualitative Response Models: A Survey,” *Journal of Economic Literature* 19, 1981, pp. 1483~1536.
10. Darmon, R. Y., “Determinants of TV Viewing,” *Journal of Advertising Research* 16, 1976, pp. 17~24.
11. Gensch, D. H. and P. Shaman, “Models of Competitive Television Ratings,” *Journal of Marketing Research* 17, 1980a, pp. 307~315.

12. \_\_\_\_\_, "Predicting TV Ratings," *Journal of Advertising Research* 20, 1980b, pp. 85~92.
13. Henry, M. D. and H. J. Rinne, "Predicting Program Shares in New Time Slots," *Journal of Advertising Research* 24, 1984, pp. 9~17.
14. Maddala, G. S., *Limited Dependent and Qualitative Variables in Econometrics*, N.Y.: Cambridge University Press, 1983.
15. McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behaviour," *Frontiers in Econometrics* by P. Zarembka (ed.), N.Y.: Academic Press, 1973.
16. \_\_\_\_\_, "Econometric Models of Probability Choice," *Structural Analysis of Discrete Data: With Econometric Applications* by C. Manski and D. McFadden (eds.), MIT Press, Cambridge, 1981.
17. Rust, R. T., *Advertising Media Models: A Practical Guide*, Lexington Books, Lexington, MA., 1986.
18. \_\_\_\_\_ and M. I. Alpert, "An Audience Flow Model of Television Viewing Choice," *Marketing Science* 3, 1984, pp. 113~124.
19. \_\_\_\_\_ and N. V. Eechambadi, "Scheduling Network Television Programs: A Heuristic Audience Flow Approach to Maximizing Audience Share," *Journal of Advertising* 18, 1989, pp. 11~18.
20. \_\_\_\_\_, Kamakura, W. A. and M. I. Alpert, "Viewer Preference Segmentation and Viewing Choice Models for Network Television," *Journal of Advertising* 21, 1992, pp. 1~18.
21. Tavakoli, M. and M. Cave, "Modelling Television Viewing Patterns," *Journal of Advertising* 25, 1996, pp. 71~86.
22. Tellis, G. J., "Advertising Exposure, Loyalty and Brand Purchase: A Two-Stage Model of Choice," *Journal of Marketing Research* 25, 1988, pp. 134~144.