

87, 92 대통령 선거 결과 예측에 대한 그래프 분석¹⁾

이 기 원²⁾, 이 승 민³⁾, 김 유 정⁴⁾

요 약

1987년도와 1992년도에 한국 꺾림에서 대통령 선거 결과를 예측하기 위하여 실시한 바 있는 여론 조사 결과를 그래픽으로 분석하였다. 1987년도에는 각 후보에 대한 직접적인 지지도보다 종교별 후보 지지도 예측에 발표의 초점이 맞춰져 있었으므로 이 발표 결과를 다각적으로 분석하였고, 1992년도에는 각 후보의 직접적인 지지도 분석에 초점을 두었다.

1. 서론

1987년도의 대통령 선거는 1971년도에 치러진 대통령 선거 이래 최초로 직접 대통령을 뽑는 선거라는 큰 의의를 가질 뿐 아니라 이승만 정권 이래 최초로 문민 정부를 수립할 수 있는 기회였다는 역사적 의미를 갖는다. 그러나 민주주의의 꽃이라고 할 수 있는 선거의 실제 시행 과정에서 그 결과를 미리 알고 싶어하는 유권자들의 욕구를 외면한 채 후보 당락에 대한 여론 조사 결과 발표를 선거법으로 금지하여 유권자들의 알 권리가 무시되고 한편으로는 통계여론조사 기법이 발달할 수 있는 좋은 기회를 놓친 결과가 되었다. 수차례에 걸친 선거법의 개정에도 불구하고 선거 기간 중의 여론조사발표는 커녕 선거 당일의 출구조사마저 현실적으로 시행하기 어려운 상태에 머무르고 있어서 각 방송사들이 합동으로 실시한 1996년 4월11일의 총선 예측에서도 예측결과가 빗나가는 촌극을 빚게 되었다.

한편 이와 같이 선거 당락에 대한 직접적인 발표가 법으로 금지되어 있었음에도 불구하고 한국 꺾림에서는 발표 자체를 종교별 후보 지지도 조사라 하여 법망을 피하면서 약간의 계산 능력만 있으면 실제 후보별 지지도를 알 수 있도록 하였다. 한국 꺾림의 발표 내용은 표 1에 수록된 바와 같으며 이는 조선일보의 의뢰로 시행되어 주간조선(1987년 12월 12일자)에 게재된 것이다. 표 1의 괄호 안 숫자들은 조사에 응답한 표본의 크기를 의미한다. 한 가지 주의할 점은 응답자 총계 2206명은 전체 유권자를 대표한다기보다는 선거 당일 실제로 투표에 참여한 유권자 전체를 대표한다고 보아야 한다. 이는 소위 target population과 sampled population의 차이에 대한 예로 들기에 적절한 경우이다. 또한 표 1에서 유의할 점은 각 종교별로 백분률을 내어 법에 저촉되는 것을 피하고 결정적인 내용, 즉 후보별 지지도는 발표하지 않았으나 약간의 계산 능력만 있으면 누구나 쉽게

1) 이 연구는 1996년도 한림대학교 학술연구조성비에 의하여 이루어졌음.

2) (200-702) 강원도 춘천시 옥천동 1번지 한림대학교 통계학과 부교수.

3) (200-702) 강원도 춘천시 옥천동 1번지 한림대학교 통계학과 부교수.

4) (200-702) 강원도 춘천시 옥천동 1번지 한림대학교 통계학과 박사과정.

표 1: 1987년도 대통령선거 종교별 후보 지지도 조사(단위 %)

	노태우	김영삼	김대중	김종필	기타
불교(631)	45.8	24.4	20.0	9.7	0.2
개신교(398)	21.1	34.9	36.4	7.3	0.3
카톨릭(152)	17.1	34.2	37.5	7.2	3.3
무(1025)	35.2	28.5	28.0	7.8	0.5
계(2206)					

판단을 할 수 있다는 것이다. 예를 들어서 노태우 후보의 지지도는

$$\frac{631 \times 45.8\% + 398 \times 21.1\% + 152 \times 17.1\% + 1025 \times 35.2\%}{631 + 398 + 152 + 1025} = 34.5\%$$

로 예측된다는 등이다.

이 자료는 또한 그 당시 유권자, 보다 정확하게 표현하자면 선거에 직접 참여한 유권자 들의 종교 분포를 말하여 준다. 한발 더 나아가 이 자료를 뒤집어 보면 각 후보별로 그를 지지한 유권자들의 종교별 분포도 알 수 있다.

본 논문에서는 이 자료와 실제 선거 결과를 막대그래프 및 원형그래프로 나타내어 시각적으로 분석하고 1992년도에 실시한 한국 갤럽의 예측과 실제 결과는 원형그래프로 나타내어 통계학 교육에 활용할 수 있게 하였다. 분석에는 S-Plus를 활용하였는데, 여기 사용된 source code는 저자들로부터 얻을 수 있다.

2. 1987대통령 선거 결과 예측 자료의 분석

이 절에서는 먼저 한국 갤럽이 발표한 내용대로 종교별 후보 지지도를 막대그래프로 나타내고 이 방법을 응용하여 각 후보의 지지도와 후보별 지지자의 종교 분포를 또 다른 막대그래프로 나타낸다. 이 작업을 보다 효과적으로 수행하기 위해서는 각 종교별 백분율이 아닌 실제의 관찰 숫자가 필요하다. 표 1에서도 절삭 오차 등으로 인하여 각 행의 합이 100 %가 아닌 경우를 여럿 관찰할 수 있다. 표 2는 표 1을 실제 관찰숫수로 바꾼 것이다.

표 2 : 1987년도 대통령선거 종교별 후보 지지도 조사(단위 명)

	노태우	김영삼	김대중	김종필	기타
불교(631)	289	154	126	61	1
개신교(398)	84	139	145	29	1
카톨릭(152)	26	53	57	11	5
무(1025)	361	292	287	80	5
계(2206)					

이 자료는 또한 당시 유권자들의 종교 분포를 추론할 수 있게 해 준다. 표 1로부터 각 종교별 응답자 수를 백분율로 나타내면 다음과 같다.

불교	개신교	카톨릭	없음
28.6%	18.0%	6.9%	46.5%

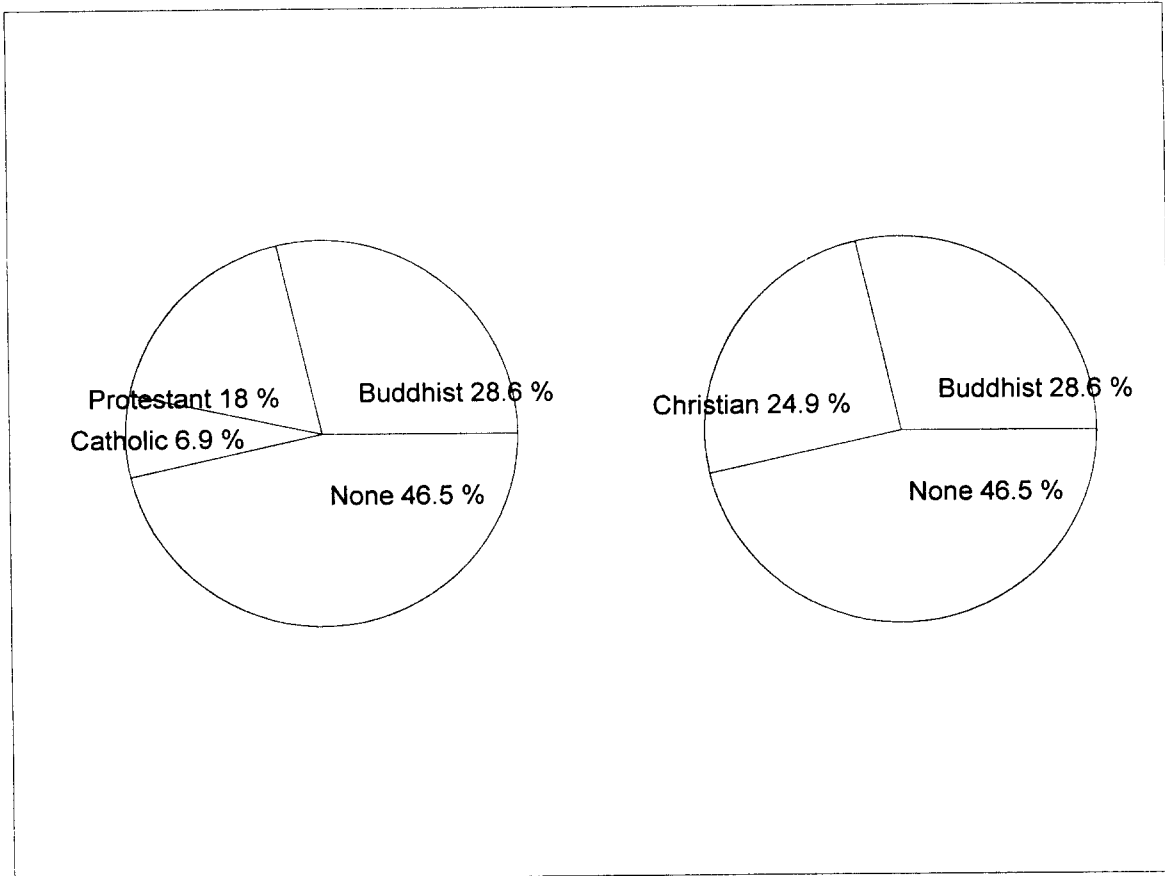


그림 1: 87대선 당시 유권자의 종교 분포.
(오른쪽 도표에서는 개신교와 카톨릭을 합쳐서 나타내었음.)

이와 같이 백분률로 된 표를 시각적으로 파악하는 가장 좋은 방법은 원형그래프를 그리는 것이다. 그림 1의 왼쪽은 위의 표를 그대로 원형그래프로 옮겨 본 것이고 오른쪽은 이 중 개신교와 카톨릭을 합친 후 원형그래프로 옮긴 것이다. 이와 같이 개신교와 카톨릭을 합쳐 보게 된 배경은 이 두 집단의 지지 성향이 몹시 닮았기 때문이다. 이 여론 조사에 사용된 표본추출방법은 3단계 집락추출 방법이다. (한국갤럽조사연구소 (1989) 참조.) 그러나 이를 마치 단순랜덤추출로 뽑은 것 처럼 가정하여 카이제곱 동일성 검정을 수행하여 보면 통계량 값은 0.8235로 계산되어 자유도 3인 카이제곱분포표로부터 p -value가 0.84임을 알 수 있다. 물론 이 계산 과정에서 기대숫수가 지나치게 작고 주요 후보의 지지도 예측에 아무런 도움을 주지 못하는 기타 후보지지는 제외하였다.

이 원형그래프로부터 시각적으로 느낄 수 있는 것은 생각보다 종교를 갖지 않은 유권자들이 거의 절반에 가까울 정도로 많이 분포하며 불교 신도가 개신교와 카톨릭 신도를 합친 것보다도 많이 분포하고 있다는 점일 것이다.

이제 종교별 후보 지지도를 막대그래프로 나타내자. 즉, 각 종교별·응답자 수를 높이로 하는 막대들 안에 각 후보들의 지지자 수를 높이로 하여 쌓아 놓은 형태의 그림을 그린다. 그림 2의 왼쪽은 표 2에서 계산한 각 종교별 후보지지자의 수효를 막대그래프로 나타낸 것이다. 그림 2의 오른쪽

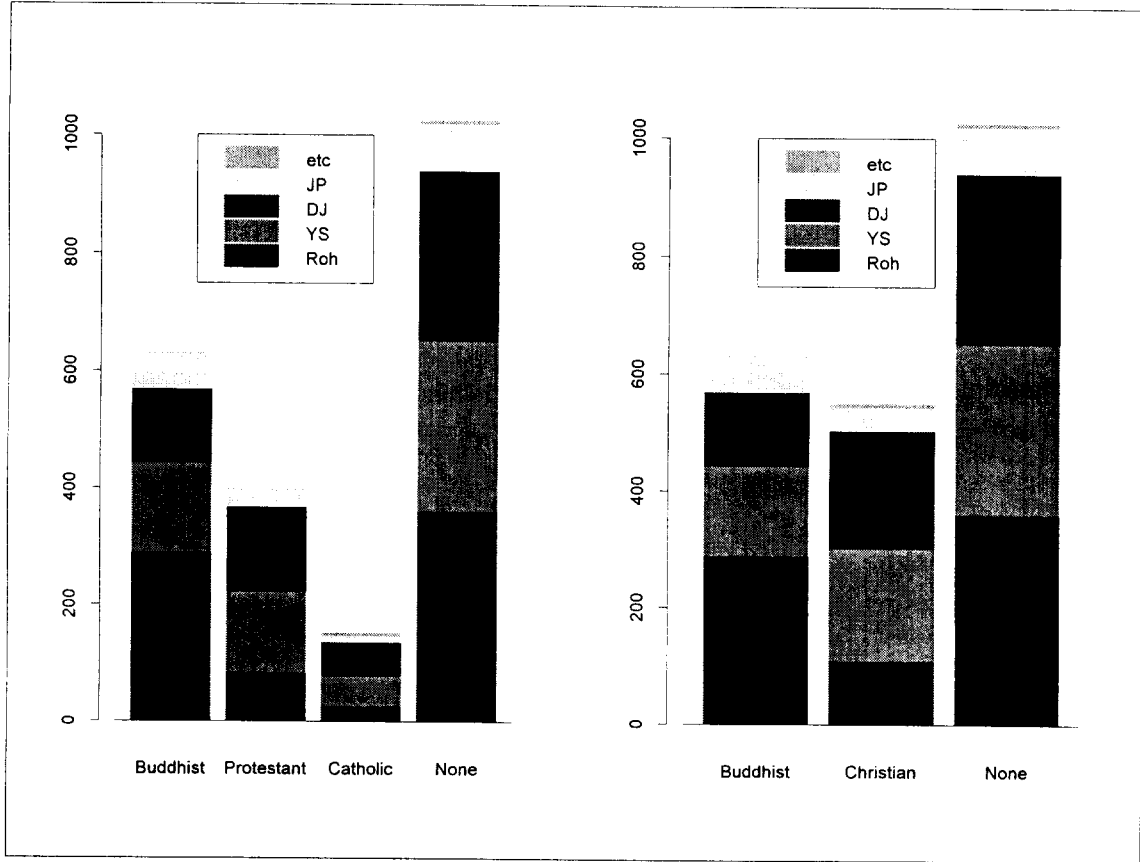


그림 2: 1987대선 종교별 후보 지지도

에는 개신교와 카톨릭을 합쳐 나타내었다. 그림 1과 비교하였을 때 원 안의 각도 크기로 표현되었던 각 종교 분포가 막대의 상대적인 높이로 나타나고 있음을 알 수 있다. 이 사실은 왼쪽에 그린 막대그래프보다 오른쪽에 개신교와 카톨릭을 합쳐 그린 막대그래프에서 보다 확실히 알 수 있다.

다음으로 이 그림에서 살필 점은 먼저 표 2에서 숫자로만 인식되었던 응답자들의 수효가 막대의 높이로 현저하게 탈바꿈하였다는 사실이다. 표 2상에서는 종교를 갖지 않은 사람들의 비율이 카톨릭 교도들보다 엄청나게 많다는 사실이 잘 느껴지지 않지만 이와 같이 막대의 높이로 나타내었을 때에는 현저한 차이가 있음을 알 수 있다.

대체적으로 개신교와 카톨릭 교도들의 투표 성향이 거의 비슷한 반면에 불교도들의 투표 성향이 노태우 후보쪽으로 많이 기울어져 있음을 알 수 있다. 이는 오른쪽 도표에서 보다 명확히 알 수 있다. 또한 그 비중이 워낙 큰 탓도 있겠지만 종교를 갖지 않은 사람들의 투표 성향이 거의 전체 투표 성향을 잘 대변하고 있는 것으로 여겨진다. 만약 표 1의 백분율을 그대로 사용하여 이 막대 그래프를 그렸다면 이와 같은 현저한 종교별 차이점이 시각적으로 명확하게 드러나지 않았을 것이다.

한편 이 자료를 뒤집어 보면 각 후보별 지지자 내에서의 종교 분포를 살펴 볼 수도 있다. 그림 3은 각 후보별 지지도에 중점을 두면서 각 후보의 종교별 지지도를 부수적으로 살펴 볼 수 있도록 해 준다. 이 그림들이 바로 유권자들이 정말로 알고 싶은 바에 대한 답을 제공한다고 볼 수 있다.

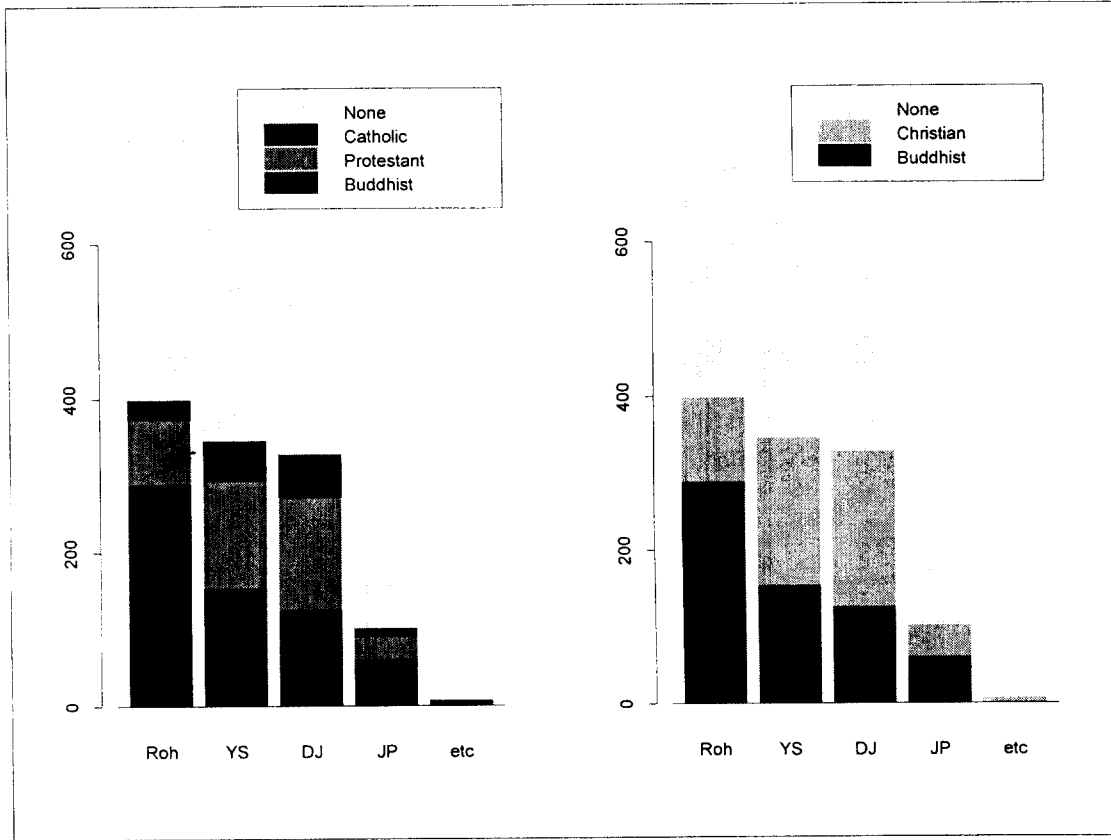


그림 3: 1987대선 후보별 지지도

물론 우리가 알고 싶은 것이 각 후보별 지지율만이라면 이러한 막대 그래프보다 그림 4의 원형 그래프가 더 간단하게 원하는 바를 보여 준다고 할 수 있다. 이 원형그래프를 그리려면 응답자 중 후보별 지지자수나 지지율을 알아야 한다. 다음 표는 서론에서 소개한 방법을 이용하여 각 후보의 지지자수를 계산하고 이를 백분율로 옮긴 것이다.

	노태우	김영삼	김대중	김종필	기타
지지자수	760	638	615	181	12
예상지지율	34.5%	28.9%	27.9%	8.2%	0.5%
실제지지율	36.6%	28.0%	27.1%	8.1%	0.2%

더우기 이 경우에는 역시 백분율로 주어지는 실제 결과와 비교할 수 있어서 통계적 기법에 의하여 수행된 여론 조사가 실제 결과를 얼마나 잘 예측하고 있는지 살펴 볼 수 있는 좋은 예가 된다.

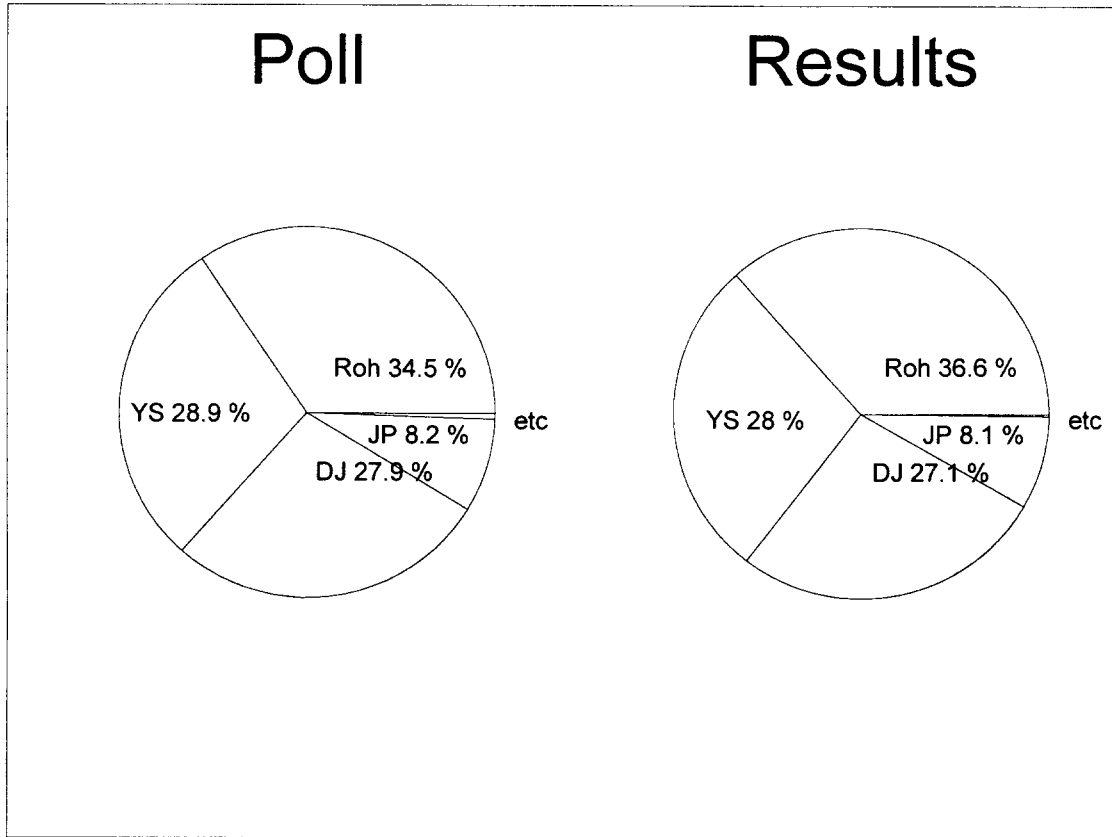


그림 4: 1987대선 후보별 지지도 : 예측과 실제 결과

3. 1992대선

이 절에서는 1992년도 대선에서 한국 갤럽의 예측과 실제 결과를 원형그래프로 비교하고자 한다. 이 분석에 사용된 자료는 다음과 같다.

	김영삼	김대중	정주영	박찬중	기타
예상지지도	39.6%	31.1%	15.7%	12.4%	1.2%
실제지지도	42.0%	33.8%	16.3%	6.4%	1.5%

다음의 그림 5에서는 1992년도 대통령선거에서 한국 갤럽이 실시한 여론 조사와 실제 결과를 원형그래프로 비교하고 있다. 박찬중 후보에 대한 지지도가 예상보다 적게 나와서 1987년도의 대통령 선거를 예측했을 때 보다는 덜 정확한 예측이었다고 할 수 있다. 이 당시 한국 갤럽이 발표한 예측오차는 2% 내외로서 이보다 훨씬 많이 차이하고 있다. 박찬중 후보를 지지하던 유권자들이 김영삼 후보에게 2.5%, 김대중 후보에게 2.7%정도 돌아선 가운데 정주영 후보에게는 상대적으

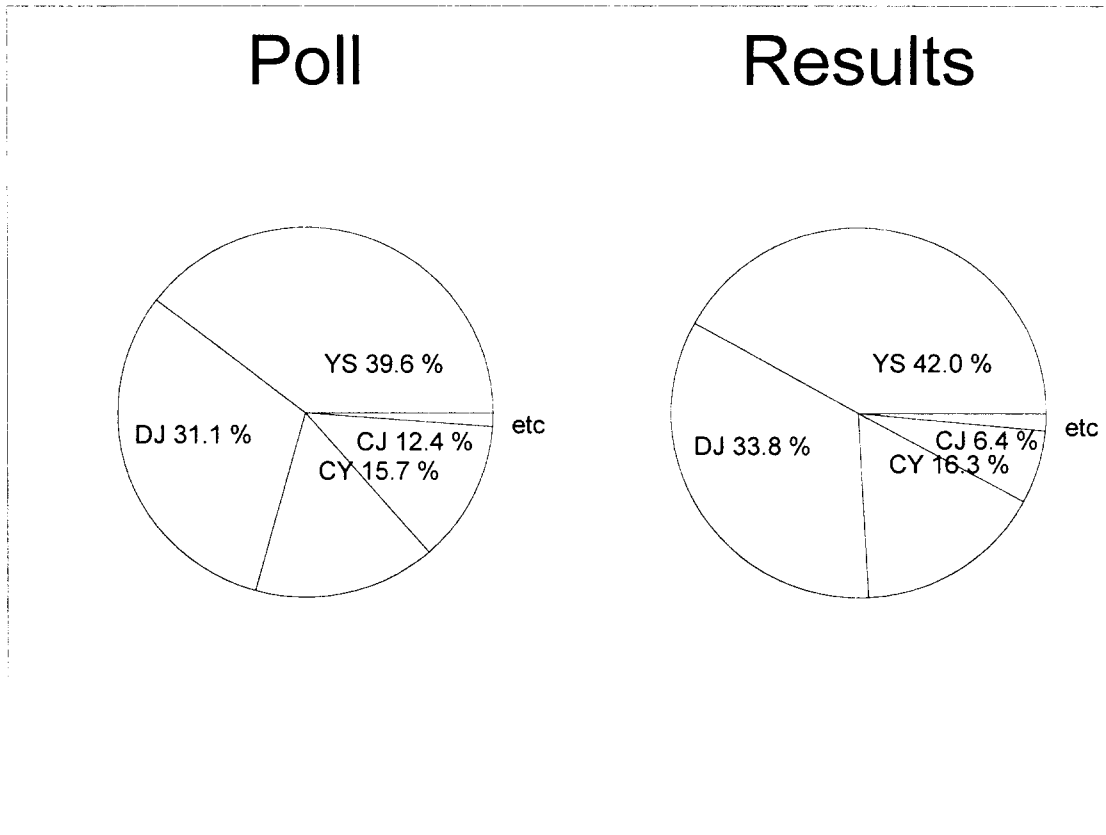


그림 5: 1992대선 후보별 지지도: 예측과 실제 결과

로 적은 0.6%정도의 지지를 보내 주어 이 당시 유권자들의 한 투표 성향을 짐작할 수 있게 해 준다.

4. 결론

1987년도와 1992년도에 한국 갤럽이 실시한 여론 조사 결과를 막대그래프와 원형그래프로 분석하였다. 1987년도 자료에서는 각 후보에 대한 지지도를 종교별로 나타내었기 때문에 각 후보별 지지도를 계산하려면 한 단계의 계산 과정을 더 거쳐야하지만 그 당시 유권자들의 종교 분포와 각 후보별 종교 분포 등의 부수적인 정보를 아는 데 많은 도움이 되었다. 이러한 부수적인 정보는 막대그래프의 각 막대 내에서의 상대적인 높이로 요약되었다. 1992년도 대통령 선거의 경우에는 1987년도 만큼 정확치는 않으나 박찬종 후보에 대한 지지도가 다른 후보들에게 몰린 결과로 판단되었다.

참고문헌

- [1] 김용옥 (1989). 「나는 불교를 이렇게 본다」, 통나무.
- [2] 한국갤럽조사연구소 (1989). 「여론 조사에 대한 한국인의 인식」, 47-54.
- [3] 한국갤럽조사연구소 (1992). 「92년 한국대통령선거 예측과 적중」.