

# 자가불소 결과\_ 단순 평균

분석 데이터: HR 인재유형 데이터 vs 역량검사 데이터

작성일: 2025년 9월 23일

분석 대상: 63명 (HR 158명, 역량검사 152명 중 매칭)

## 1. 핵심 요약

- 데이터 매칭: HR 158명, 역량검사 152명 중 최종 매칭 63명(39.9%)
- 핵심 역량: 전략성(24.8%)·정체성(14.9%)·사회성(14.5%)이 인재 선발에 가장 중요한 요인으로 도출
- 선발 모델: Correlation 기반 가중치가 가장 우수한 성능 ( $F1=0.744$ ,  $AUC=0.752$ )
- 통계적 유의성: 퍼뮤테이션 테스트로 모델의 예측력 검증 ( $p<0.01$ )

## 2. 데이터 로딩 및 전처리

### 2.1 전체 데이터 로딩 현황

#### 【STEP 1】 데이터 로딩

- HR 데이터: 158명 로드 완료
- 역량검사 데이터: 152명 (상위항목 & 종합점수 모두 확보)

#### 【STEP 1-1】 신뢰가능성 필터링

- 역량검사 데이터 신뢰불가: 5명 제외
- 최종 신뢰가능 데이터: 147명

#### 【STEP 1-2】 개발자 여부 필터링

- 개발자: 29명 (19.7%)
- 비개발자: 118명 (80.3%)
- 분석 선택: 모든 데이터 사용 → **147명 유지**

### 최종 매칭 결과

- 최종 매칭 인원: 63명
- 매칭률: 39.9%

### 2.2 인재유형 분포

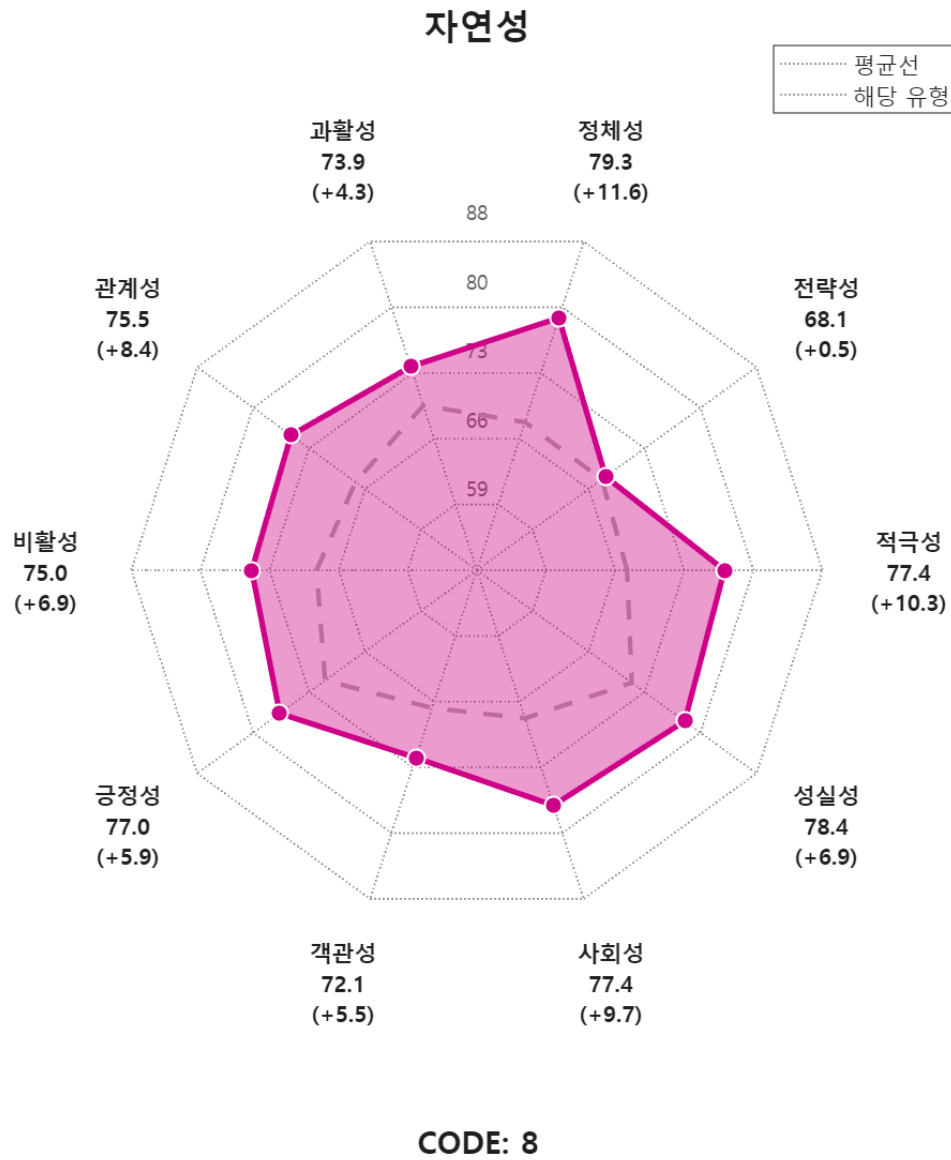
| 인재유형    | 인원(명) | 비율(%) |
|---------|-------|-------|
| 성실한 가연성 | 28    | 44.4  |
| 유능한 불연성 | 10    | 15.9  |
| 게으른 가연성 | 6     | 9.5   |
| 소화성     | 6     | 9.5   |
| 무능한 불연성 | 5     | 7.9   |
| 유익한 불연성 | 4     | 6.3   |
| 자연성     | 4     | 6.3   |

특징:

- 성실한 가연성이 전체의 44.4%를 차지하는 편중된 분포

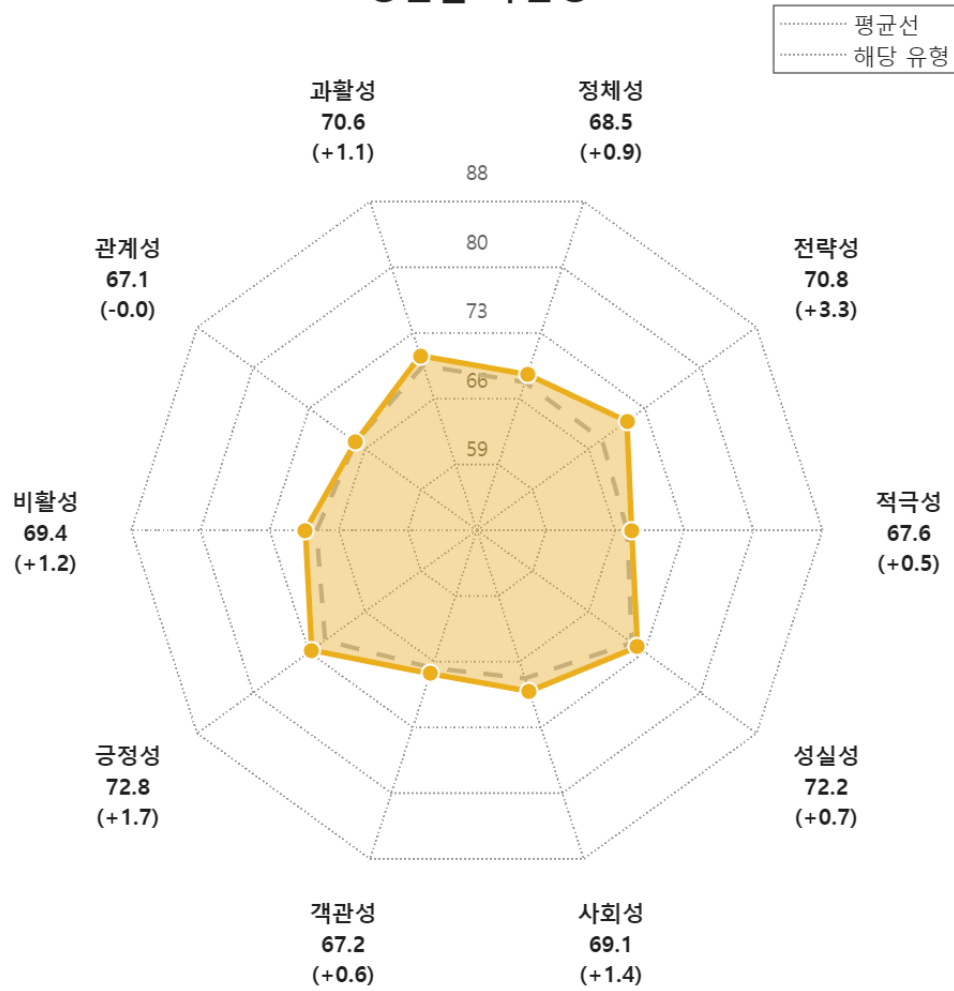
| 유형 (CODE)   | 비활성         | 정체성          | 적극성          | 전략성          | 성실성          | 사회성         | 관계성         |
|-------------|-------------|--------------|--------------|--------------|--------------|-------------|-------------|
| 자연성 (8)     | 75.0 (+6.9) | 79.3 (+11.6) | 77.4 (+10.3) | 68.1 (+0.5)  | 78.4 (+6.9)  | 77.4 (+9.7) | 75.5 (+8.4) |
| 성실한 가연성 (7) | 69.4 (+1.2) | 68.5 (+0.9)  | 67.6 (+0.5)  | 70.8 (+3.3)  | 72.2 (+0.7)  | 69.1 (+1.4) | 67.1 (+0.4) |
| 유익한 불연성 (6) | 76.3 (+8.1) | 74.7 (+7.0)  | 78.5 (+11.4) | 73.0 (+5.5)  | 79.1 (+7.6)  | 73.8 (+6.0) | 73.3 (+6.0) |
| 유능한 불연성 (5) | 60.7 (-7.4) | 65.1 (-2.5)  | 60.2 (-6.9)  | 66.8 (-0.8)  | 67.7 (-3.8)  | 63.1 (-4.6) | 64.1 (-3.4) |
| 게으른 가연성 (4) | 66.6 (-1.6) | 66.7 (-1.0)  | 65.9 (-1.2)  | 64.8 (-2.8)  | 71.1 (-0.3)  | 64.2 (-3.5) | 65.0 (-2.1) |
| 무능한 불연성 (3) | 62.8 (-5.4) | 56.2 (-11.5) | 62.7 (-4.4)  | 62.2 (-5.3)  | 59.7 (-11.7) | 59.0 (-8.7) | 64.6 (-2.1) |
| 소화성 (1)     | 71.2 (+3.0) | 66.0 (-1.7)  | 66.8 (-0.3)  | 56.9 (-10.6) | 74.7 (+3.2)  | 68.9 (+1.2) | 66.5 (-0.1) |

### 3.1. 자연성 (CODE: 8)



### 3.2. 성실한 가연성 (CODE: 7)

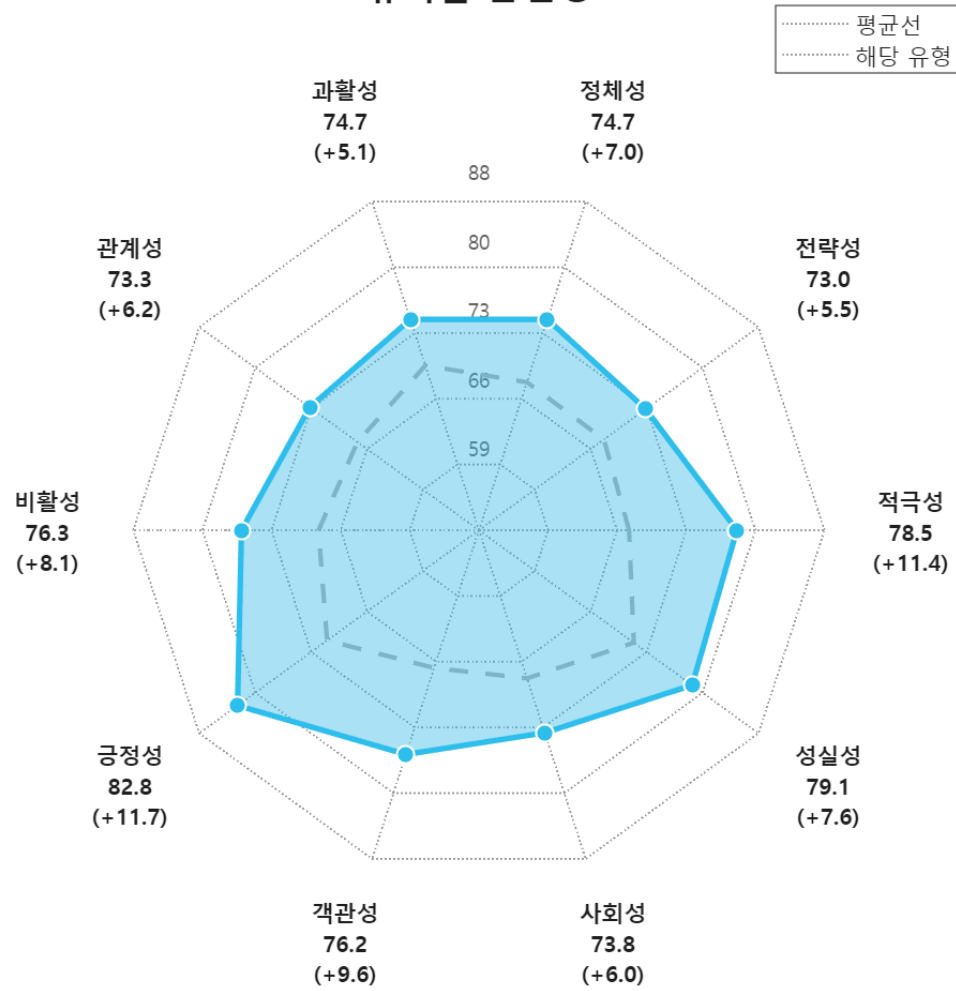
## 성실한 가연성



CODE: 7

### 3.3. 유익한 불연성 (CODE: 6)

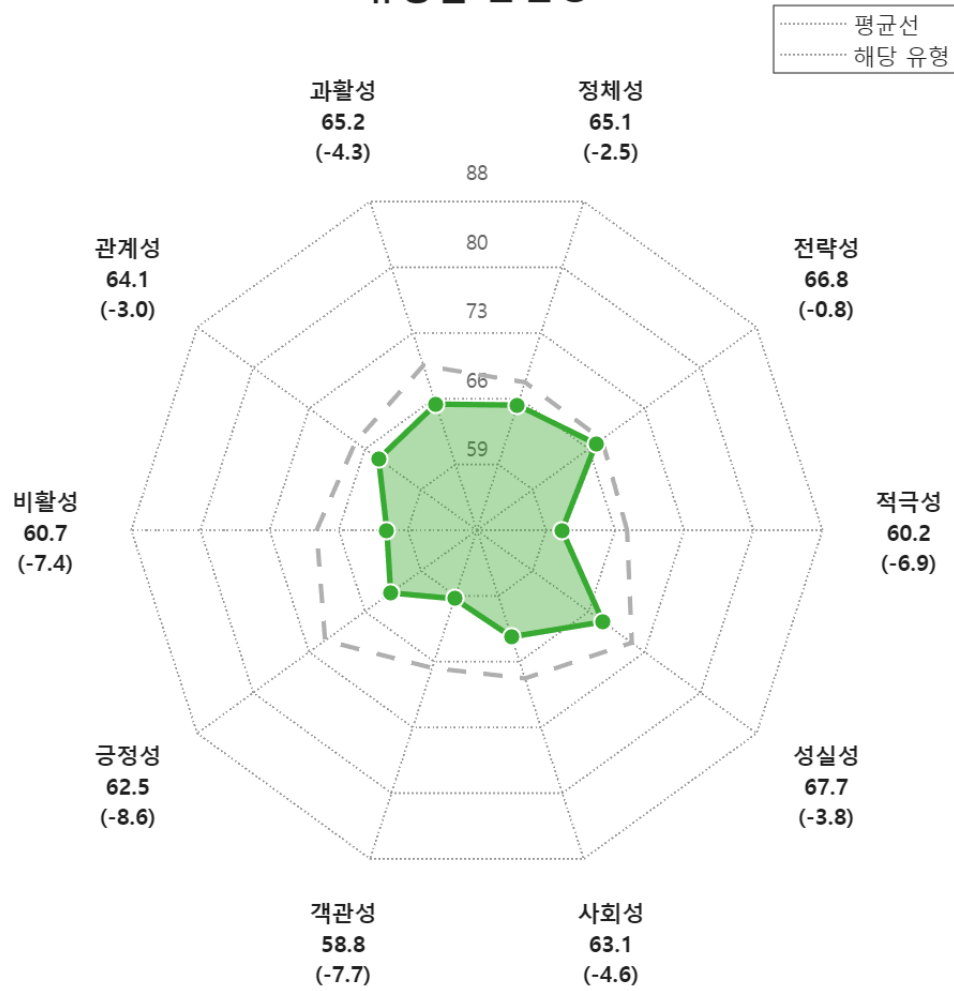
## 유익한 불연성



CODE: 6

### 3.4. 유능한 불연성 (CODE: 5)

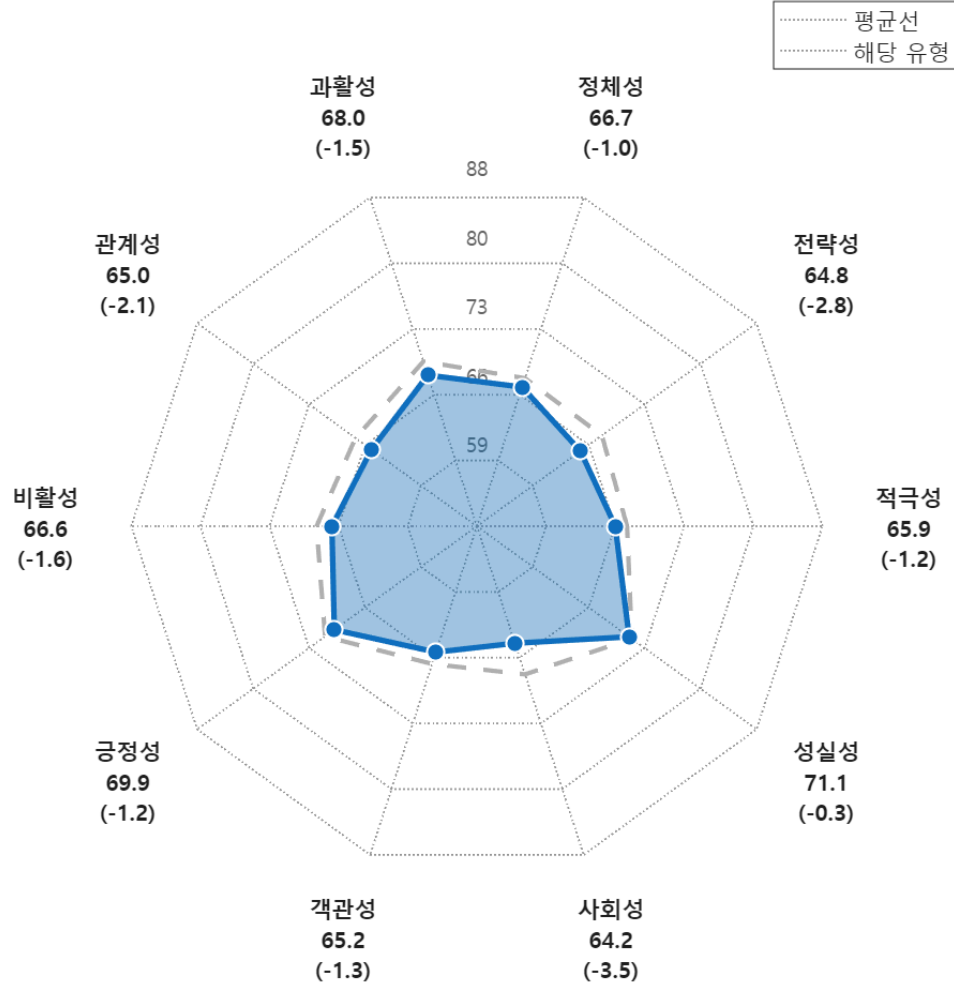
## 유능한 불연성



CODE: 5

### 3.5. 게으른 가연성 (CODE: 4)

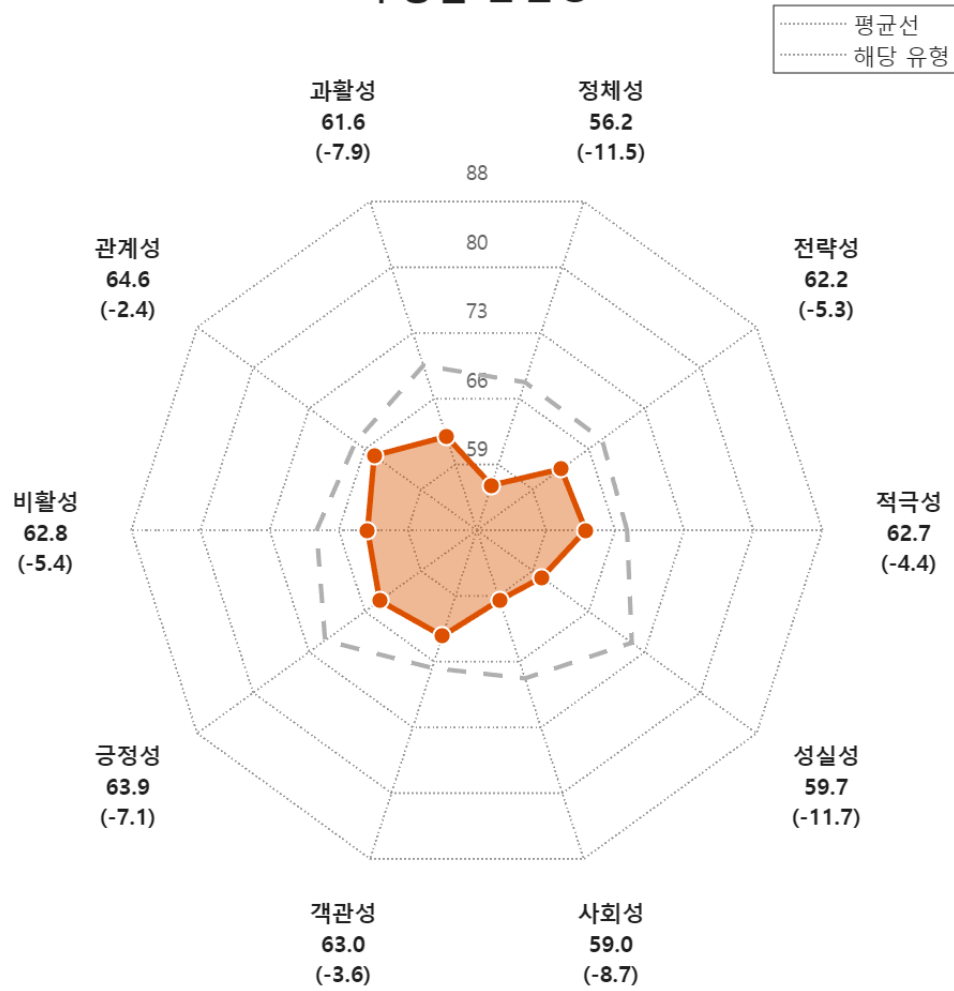
## 게으른 가연성



CODE: 4

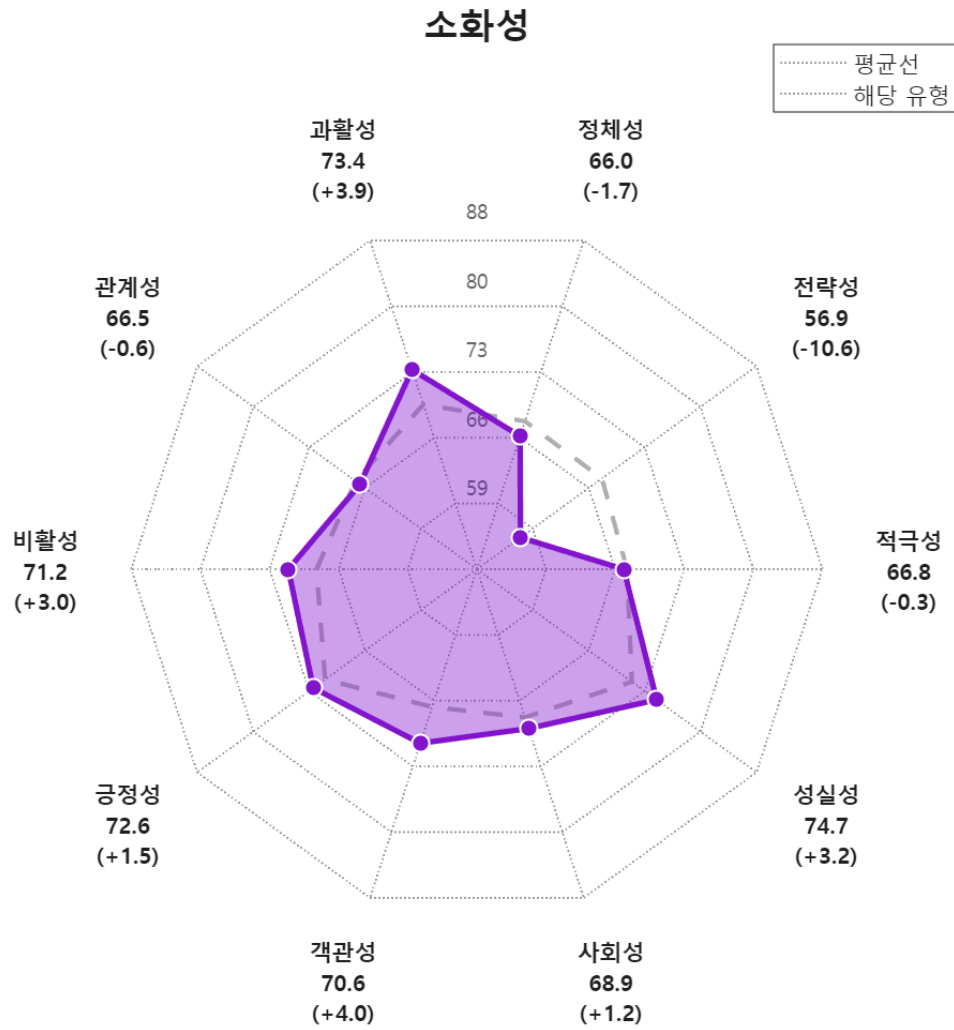
### 3.6. 무능한 불연성 (CODE:2)

## 무능한 불연성



CODE: 3

### 3.7. 소화성 (CODE: 1)

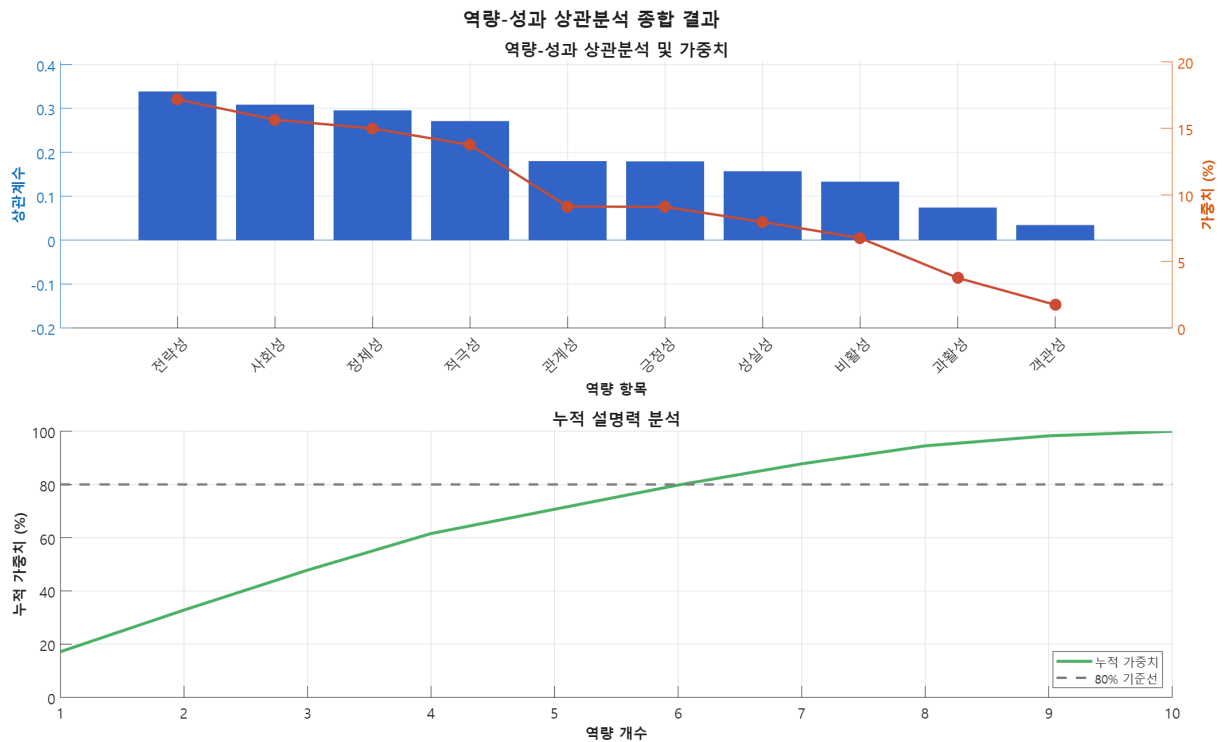


**CODE: 1**

### 3. 역량-성과 상관분석

#### 3.1 상위 10개 성과 예측 역량





| 순위 | 역량  | 상관계수   | p-값   | 효과크기 | 가중치(%) |
|----|-----|--------|-------|------|--------|
| 1  | 전략성 | 0.338* | 0.023 | 0.78 | 17.17  |
| 2  | 사회성 | 0.308* | 0.028 | 0.58 | 15.65  |
| 3  | 정체성 | 0.295* | 0.035 | 0.58 | 14.99  |
| 4  | 적극성 | 0.271  | 0.054 | 0.31 | 13.75  |
| 5  | 관계성 | 0.180  | 0.207 | 0.25 | 9.13   |
| 6  | 긍정성 | 0.179  | 0.208 | 0.46 | 9.10   |
| 7  | 성실성 | 0.157  | 0.272 | 0.44 | 7.96   |
| 8  | 비활성 | 0.133  | 0.352 | 0.33 | 6.75   |
| 9  | 과활성 | 0.074  | 0.605 | 0.30 | 3.76   |
| 10 | 객관성 | 0.034  | 0.811 | 0.15 | 1.74   |

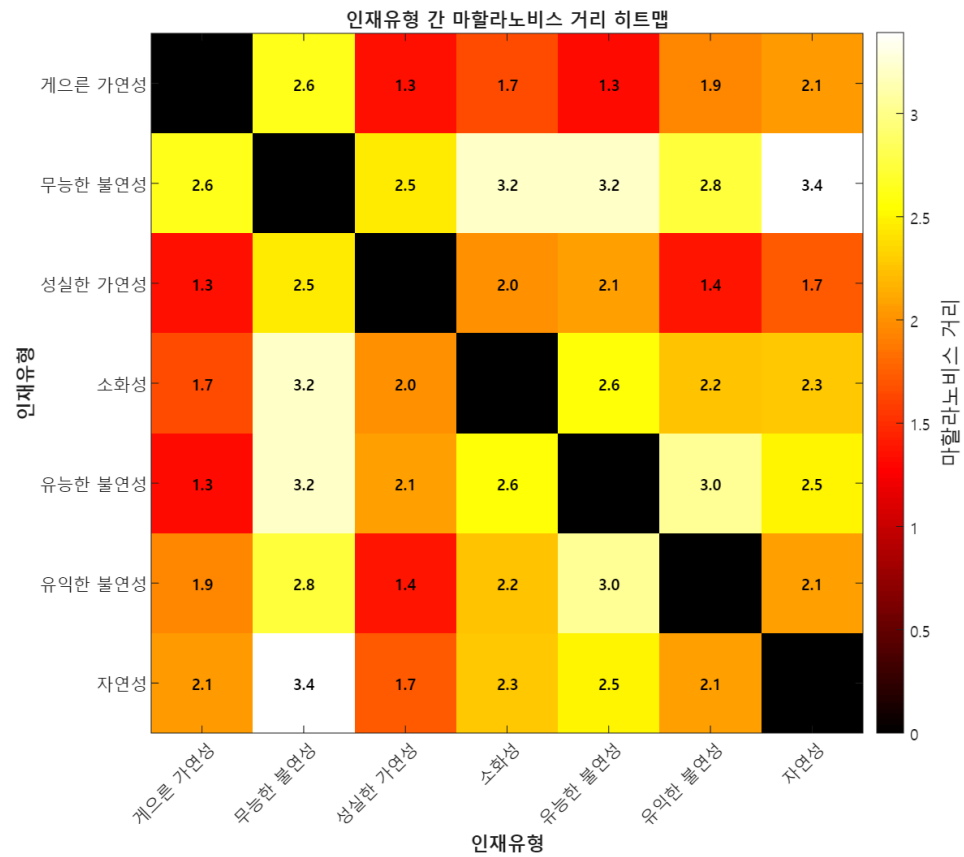
- $p < 0.05$

## 4. 인재유형 간 마할라노비스 거리 분석

### 4.1 분석 개요

- 목적: 7개 인재유형 간 유사성 정량화
- 방법: 10개 역량 기반 다변량 거리 측정

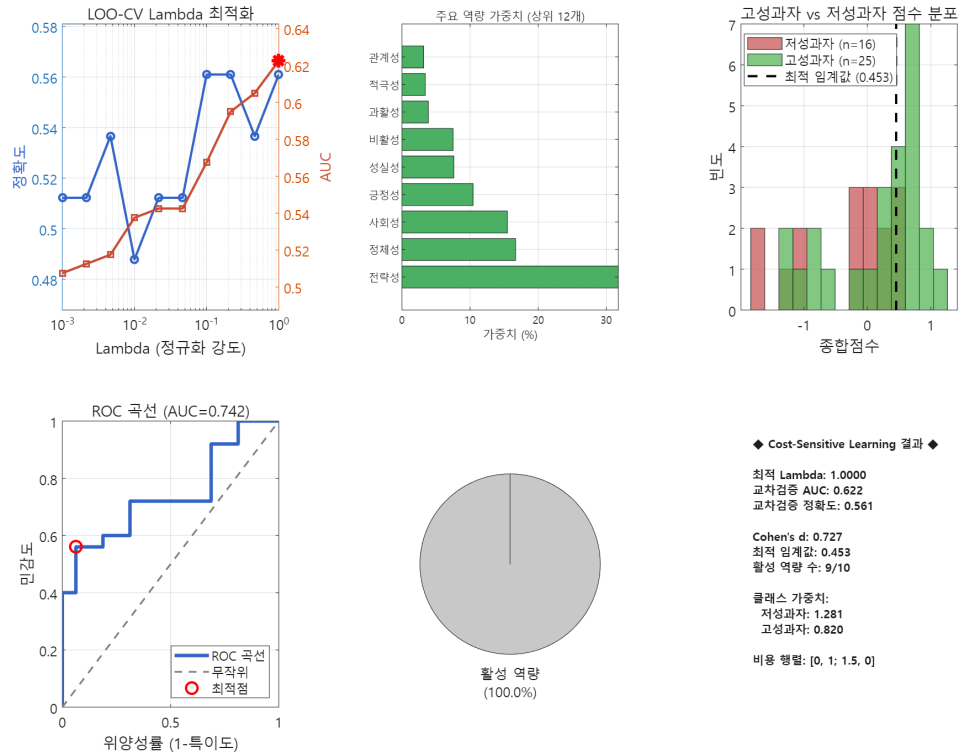
### 4.2 주요 발견



| 가장 유사한 유형 쌍       | 거리   | 가장 다른 유형 쌍    | 거리   |
|-------------------|------|---------------|------|
| 게으른 가연성 ↔ 유능한 불연성 | 1.33 | 무능한 불연성 ↔ 자연성 | 3.39 |
| 성실한 가연성 ↔ 게으른 가연성 | 1.34 | 무능한 불연성 ↔ 소화성 | 3.21 |

## 5. Cost-Sensitive Logistic Regression

## Cost-Sensitive Learning 기반 고성과자 예측 시스템 분석 결과



### 5.1 분석 설계

- 고성과자: 자연성, 성실한 가연성, 유익한 불연성 (25명)
- 저성과자: 무능한 불연성, 소화성, 게으른 가연성 (16명)
- 클래스 불균형 비율: 1.56:1
- 비용 행렬: 저성과자→고성과자 오분류 비용 1.5배 적용

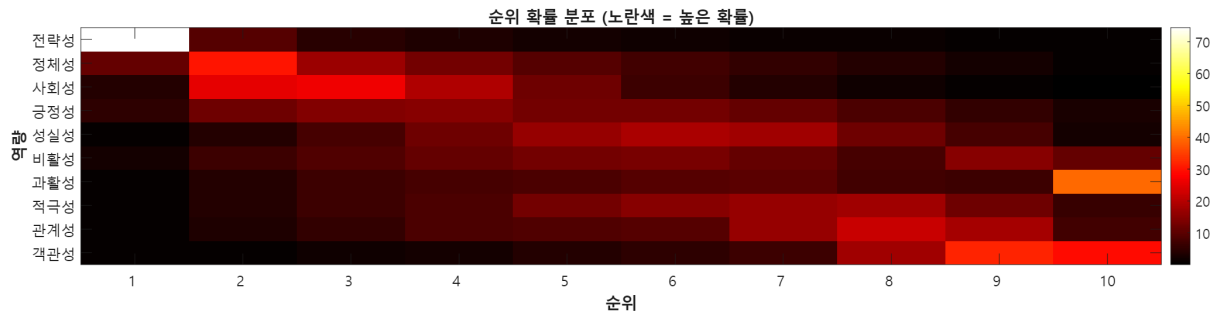
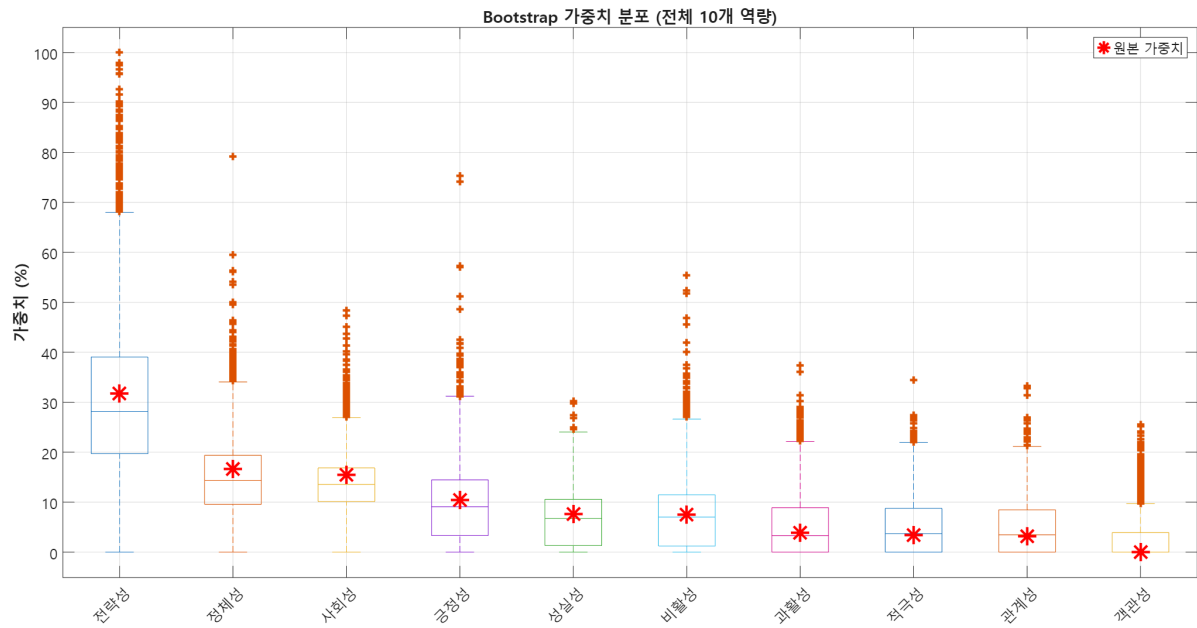
### 5.2 모델 성능

- 최적  $\lambda$ : 1.000 (Ridge 정규화)
- AUC: 0.742
- 최적 임계값: 0.453
- Cohen's d: 0.727 (큰 효과)

### 5.3 주요 역량 기여도

| 순위 | 역량  | 가중치(%) | 원계수    |
|----|-----|--------|--------|
| 1  | 전략성 | 31.75  | 0.1293 |
| 2  | 정체성 | 16.69  | 0.0680 |
| 3  | 사회성 | 15.48  | 0.0631 |
| 4  | 긍정성 | 10.44  | 0.0425 |
| 5  | 성실성 | 7.61   | 0.0310 |
| 6  | 비활성 | 7.51   | 0.0306 |
| 7  | 과활성 | 3.89   | 0.0158 |
| 8  | 적극성 | 3.43   | 0.0140 |
| 9  | 관계성 | 3.20   | 0.0130 |

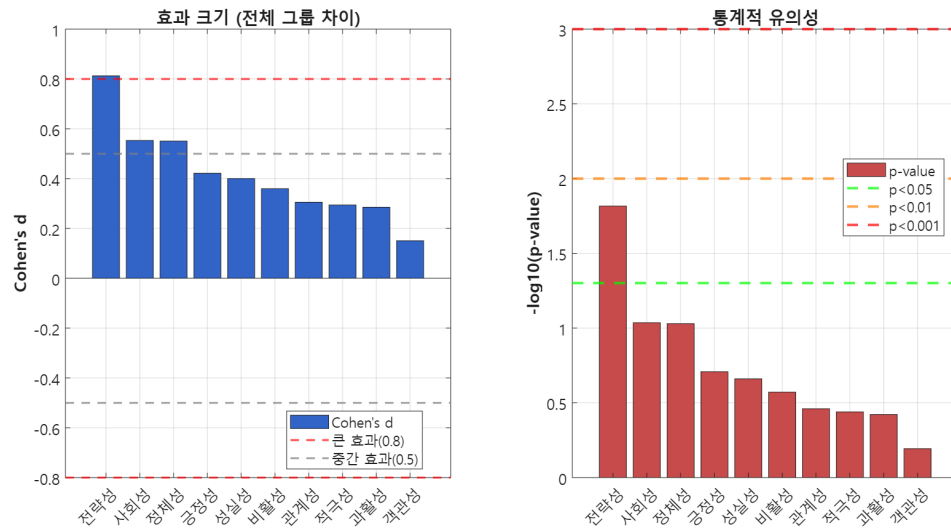
### 5.4 Bootstrap 안정성 검증 (5000회)



| 역량  | 원본(%) | 평균(%) | 95% CI       | CV   | Top5확률(%) |
|-----|-------|-------|--------------|------|-----------|
| 전략성 | 31.75 | 30.63 | [5.22-68.83] | 0.52 | 94.6      |
| 정체성 | 16.69 | 14.66 | [0.00-32.02] | 0.56 | 80.4      |
| 사회성 | 15.48 | 13.51 | [0.00-25.62] | 0.44 | 87.0      |
| 긍정성 | 10.44 | 9.66  | [0.00-26.76] | 0.81 | 59.0      |
| 성실성 | 7.61  | 6.69  | [0.00-17.28] | 0.80 | 41.3      |

## 6. 극단 그룹 t-test 비교

전체 그룹 t-test 분석 결과



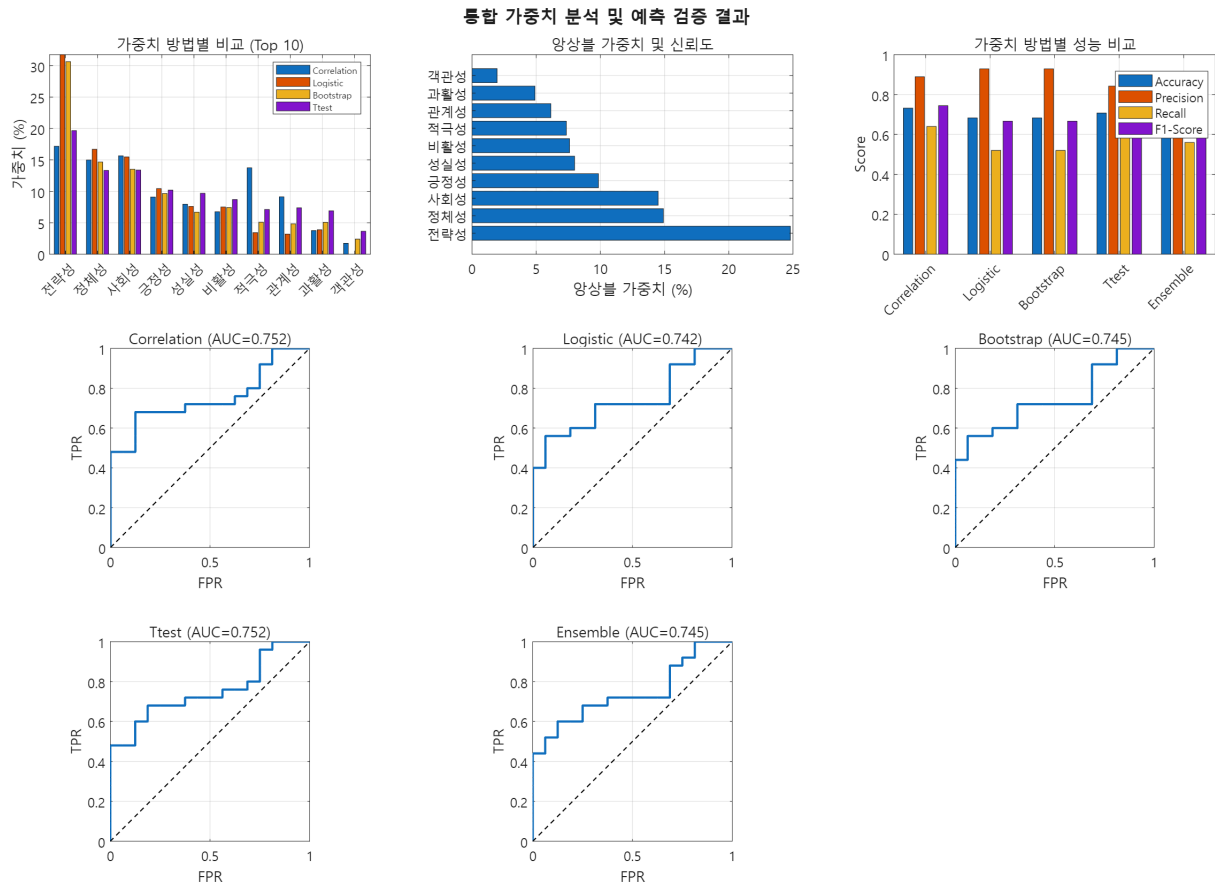
## 6.1 전체 그룹 비교 (n=41)

| 역량  | 고성과(M±SD) | 저성과(M±SD) | 차이   | t값   | p값     | Cohen's d |
|-----|-----------|-----------|------|------|--------|-----------|
| 전략성 | 71.0±12.3 | 61.5±10.4 | +9.4 | 2.54 | 0.015* | 0.812     |
| 사회성 | 70.2±10.5 | 64.1±12.1 | +6.2 | 1.73 | 0.092† | 0.553     |
| 정체성 | 69.9±11.8 | 63.2±12.8 | +6.7 | 1.72 | 0.093† | 0.551     |
| 긍정성 | 73.1±10.0 | 68.9±10.3 | +4.3 | 1.32 | 0.196  | 0.421     |
| 성실성 | 72.9±8.5  | 68.7±13.0 | +4.2 | 1.25 | 0.219  | 0.400     |

- $p < 0.05$ ,  $†p < 0.10$

핵심 차별화 역량: 전략성 ( $p < 0.05$ ,  $d > 0.8$ )

## 7. 통합 가중치 분석



#### 7.1 4가지 방법론 비교

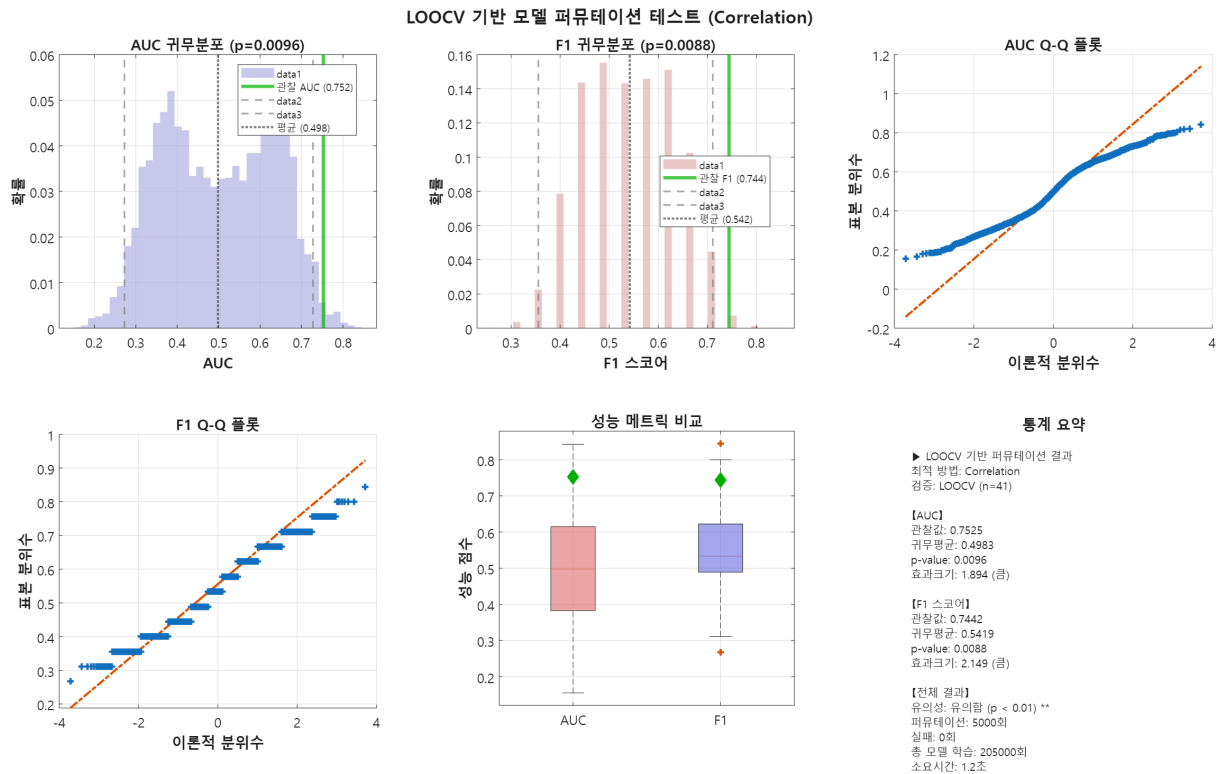
| 역량  | Corr(%) | Logit(%) | Boot(%) | Ttest(%) | Ensemble(%) | 신뢰도    |
|-----|---------|----------|---------|----------|-------------|--------|
| 전략성 | 17.17   | 31.75    | 30.63   | 19.66    | 24.80       | Medium |
| 정체성 | 14.99   | 16.69    | 14.66   | 13.33    | 14.91       | High   |
| 사회성 | 15.65   | 15.48    | 13.51   | 13.38    | 14.50       | High   |
| 긍정성 | 9.10    | 10.44    | 9.66    | 10.20    | 9.85        | High   |
| 성실성 | 7.96    | 7.61     | 6.69    | 9.69     | 7.99        | High   |

#### 7.2 방법별 예측 성능

| 방법                 | Accuracy     | AUC          | F1           | Precision | Recall |
|--------------------|--------------|--------------|--------------|-----------|--------|
| <b>Correlation</b> | <b>0.732</b> | <b>0.752</b> | <b>0.744</b> | 0.889     | 0.640  |
| Ttest              | 0.707        | 0.752        | 0.727        | 0.842     | 0.640  |
| Ensemble           | 0.683        | 0.745        | 0.683        | 0.875     | 0.560  |
| Logistic           | 0.683        | 0.742        | 0.667        | 0.929     | 0.520  |
| Bootstrap          | 0.683        | 0.745        | 0.667        | 0.929     | 0.520  |

최적 방법: Correlation (F1=0.744)

### 8. 퍼뮤테이션 테스트 결과



## 8.1 통계적 유의성 검증 (5000회)

| 방법          | 관찰 AUC | AUC p-값 | 관찰 F1 | F1 p-값  | 유의성 |
|-------------|--------|---------|-------|---------|-----|
| Correlation | 0.753  | 0.012*  | 0.744 | 0.010** | **  |
| Logistic    | 0.743  | 0.004** | 0.667 | 0.282   | **  |
| Bootstrap   | 0.745  | 0.010** | 0.667 | 0.030*  | **  |
| Ttest       | 0.753  | 0.013*  | 0.727 | 0.011*  | *   |
| Ensemble    | 0.745  | 0.014*  | 0.683 | 0.047*  | *   |

- \*p < 0.01, \*p < 0.05

## 8.2 LOOCV 기반 검증

- 샘플 수: 41명
- AUC: 0.753 (p=0.0096\*\*)
- F1: 0.744 (p=0.0088\*\*)
- 효과 크기: AUC=1.894, F1=2.149 (모두 큰 효과)

## 9. 핵심 발견 및 권장사항

### 9.1 핵심 역량 (신뢰도 High & 가중치 >5%)

- 전략성: 24.80% (일관된 1순위)
- 정체성: 14.91% (표준편차: 1.38)
- 사회성: 14.50% (표준편차: 1.22)
- 긍정성: 9.85% (표준편차: 0.60)
- 성실성: 7.99% (표준편차: 1.25)

### 9.2 오분류 패턴 분석

- 전체 오분류율: 26.8% (11/41)

- **False Positive:** 2건 (모두 게으른 가연성)
- **False Negative:** 9건 (주로 성실한 가연성)

### 9.3 권장사항

1. **예측 모델:** Correlation 방법을 주요 모델로 사용
2. **임계값:** 0.378 (민감도와 특이도 균형)
3. **추가 검토사항:**
  - 게으른 가연성과 성실한 가연성 구분 기준 재검토
  - 성별 효과 존재 ( $p=0.037$ ) - 공정성 검토 필요
  - 데이터 추가 확보 시 모델 재학습 권장

### 9.4 제한사항

- 샘플 수 제한 ( $n=41$ )으로 일반화 주의
- 클래스 불균형 존재 (1.56:1)
- Range Restriction 효과로 실제 예측력 과소평가 가능성

---

**분석 완료:** 2025년 9월 23일 18:57:07