

2022 정시 표본 분석 방법

칸 색을 정하는 기준

파란색(등록)

1. 다른 군이 모두 탈락으로 뜨고, 분석하고 있는 학과와 급간 차이도 얼마 안 나거나 (1~2 급간) 학과 차이도 별로 없을 때 (예시: 가군 건국대 컴공 탈락,
2. 다른 대학이 합격권으로 뜨지만 학과 조합이나 급간 차이를 봤을 때 분석하고 있는 학교가 다른 학교보다 더 좋아 보일 때 (예시: 다군 홍익대 컴공, 나군 과기대 컴공, 가군 부산대 컴공을 썼을 때는 먼저 나, 다군에 모두 수도권 대학을 썼다는 점에서 지방보다는 수도권에 사는 표본이라는 것을 알 수 있으며, 선호도 자체가 홍대가 과기대, 부산대보다 높기 때문에 홍대에 갈 것이라고 예상할 수 있다.)

빨간색(다른 학교 등록)

1. 분석하고 있는 학교보다 다른 학교 급간이 훨씬 더 높는데, 그 학교도 합격권으로 뜰 때 (예시: 다군 홍익대 컴공 최조합, 나군 경희대 소프트웨어 추합이면 집안 형편이 어려워서 울며 겨자먹기로 홍대를 갈 수도 있겠지만 그런, 내가 알 수 없는 것까지는 공예질을 할 수 없으니 여기까지는 따지지 않는다. 나는 최대한 '합리적 의심'을 하는 것이다. '킹리적 갓심')

회색(허수)

1. 다른 학교를 지원할 수 있는데도 다른 학과에 지원하지 않았을 때 (예시: 가, 나, 다군 중 한두 군이 비워져 있을 때)
2. 분석하고 있는 학과와 세 급간 이상 차이가 날 때 (예시: 다군 동국대, 나군 서강대, 가군 연세대 or 가군 한의대, 나군 시립대, 다군 동국대 - 정시는 대부분 자기 성적을 보고 라인을 잡은 다음에 비슷한 급간으로 쓴다.)
3. (업데이트 관찰) 표본이 움직이지 않을 때: 몇 번 모의지원을 돌려보다가 포기한 것으로 본다. 진학사와 메가 모의지원을 모두 결제했는데 표본은 메가보다 진학사에 더 많을 것이기 때문에 허수로 봐야 한다.

보라색/남색(50% 확률로 등록)

1. 급간도 비슷하고 학과도 비슷한데 두 학과에 모두 붙었을 때: 두 학교 중에서 어디로 갈지 모르기 때문에 50% 확률로 놓는다. (보라색)
2. 다른 두 군 중에서 한 학과는 뜬금없는 학과를 넣었는데 다른 한 학과는 분석하는 학과와 급간도 비슷하고 학과도 비슷한데, 다른 학과가 모두 탈락으로 뜰 때: 허수일 확률이 50%이기 때문에 남색으로 칠한다.

이때는 아래와 같이 분석한다.

보라색/남색 표본 분석법

아래 표본 예측 결과에서, 남색이나 보라색 칸은 이 학과에 넣을 확률이 1/2 라는 것을 뜻하며, 내 위에 있는 인원 수를 확률변수 X 로 잡고 아래와 같은 확률분포를 세운다.

$$X \sim B\left(n, \frac{1}{2}\right)$$

여기서 n 은 아래 표에서 남색이나 보라색 칸의 수다. 확률을 실제로 계산할 때는 보라색이나 남색 칸이 충분히 많을 것이기 때문에 아래 정규분포로 계산한다. 확률은 그냥 계산하기 편하라고 $\frac{1}{2}$ 로 잡았다.

$$X \sim N\left(\frac{1}{2}n, \frac{1}{4}n\right)$$

그러면 내 위에 있을 사람들은 이 표에서 파란색으로 채워질 것인데, 보라색이나 남색은 파란색으로도 채워질 수 있고 빨간색이나 회색으로도 칠해질 수 있기 때문에 아래 확률을 정규분포를 표준화해서 계산하면 된다. 여기서 확률변수 X 는 이미 파란 칸이 아니라 보라색, 남색 칸 중에서 파란색으로 칠해질 칸의 수이고, k 는 내가 합격하기 위해 내 위에 있을 수 있는 최대 표본 수 m 은 이미 파란색으로 칠해진 칸의 수다. 쉽게 말해서 내 위에 k 명 아래로 있어야 내가 이 학과에 붙는다는 것이다.

$$P(X \leq k - m)$$

실제로 엑셀에서 계산을 할 때는 아래 함수를 이용한다.

`=BINOM.DIST(점공상 합격인원-파란 칸 수, 보라색+남색 칸 수, 1/2, TRUE)`

여기서 점공상 합격인원은 합격예측 리포트에서 최초합으로 뜨는 마지막 등수와 같다.