

<표본분석에 대하여>

우선, 분석하는 방법은 다양합니다. 왜냐하면, 미분법처럼 정해진게 아닙니다. 다항함수를 미분 하라하면 미분법에따라 해야겠죠. 근데 표본은 따로 어떤식으로 분석해야만한다는 그런게 없습니다. 그러니, 이런 분석방법도 있구나 하면 됩니다. 누군가는 정말 무식하게 하는구나 싶기도 할거고, 누군가는 이럴 수도 있겠거니 할겁니다. 상관없습니다. 저는 이런식으로해서 결국엔 누차하지만 스나로 들어갔으니깐요.

표본분석의 절차는 크게

- 1.업데이트 될 때마다 표본을 수집한다
- 2.표본 수집후 해석한다
- 3.1과 2를 누적해서 반복한다
- 4.누적된 데이터와 경쟁률을 고려하여 원서를 쓴다

...네 이게 답니다. 그리고 당연히 진학4 표본분석을 말합니다. 그리고 표본해석은 반드시 엑셀없으면 못합니다. 엑셀함수를 적어도 쓸 줄은 알아야합니다.

1.업데이트 될 때마다 표본을 수집한다

1)표본 수집의 범위

우선, 원칙적으로 s대부터 본인이 원하는 라인까지 수집해야합니다. 만약에 본인이 서성한쓰신다면 s대부터 ky, 서성한까지고, 중경외시쓰면 중경외시전라인이랑 s부터 모조리 수집해야 합니다. 정말 어렵다면 s대는 빼더라도 ky부터 분석해야합니다.

==>정말 어렵더라면, 자기 윗라인이라도 분석해야합니다. 왜냐하면, 윗라인사정에 따라서 아랫라인에 영향을 받습니다. s대는 윗라인이 없으니 차치하더라도, 당장 ky는 s대 뺄쪽에 따라서 ky 빵이 결정되고, 마찬가지로 ky쪽뺄에 따라서 가군 서성(한) 빵이결정되고, 가군 서성한 뺄쪽에 따라서 나군 성한 뺄쪽이 결정되고... 또 ky에 얼마나 몰렸느냐에 나군 성한 빵이 결정되고... 나군 성한뺄쪽과 교대에 따라서, 가군 중경시 빵이 결정되고... 이런식으로 계속갑니다. 만약 본인이 중경시라인이라 하더라도 결국엔 위부터 바라보는게 가장 스나가 정확해집니다.

그리고 같은 과에선 모든 과를 찾으셔야합니다 한편 본인이 중경외시라인을 주력으로 분석한다면 경희대 경우 한의대나 이대의 의예 이런 과는 제외하셔도 좋습니다.

2)표본 수집의 대상

데이터는 여러 가지를 담습니다. 예를 들어서 저라는 데이터는 남자이고 나이는 몇 살이고... 여자친구는 있고 없고... 근데 원서입장에서 무슨 쓸모가 있나요? 결국 원하는 데이터는 성적 이 그 대학에서 그 학과에서 몇등이고 어떻게 원서를 쓸지가 중요하죠.

그래서 저는 다음과 같이 표본의 '데이터'를 수집했습니다. 각 해당과에서

- (1)나의등수 - 실지원 20개를 했을 때 자신의 표본상 등수를 의미합니다.
- (2)실지원자수 - 모의지원상 실지원자 총수를 의미합니다.
- (3)진학4에서 제시하는 각과의 예상경쟁률 - 그 말대로 진학4가 제시하는 경쟁률을 의미합니다.
- (4)진학4에서 제시하는 최초합격인원자수 - 모의지원 상세보기에서 진학4에서 제시하는 최초

합격자수를 의미합니다. 직접 세셔야합니다.

(5)진학4에서 제시하는 추가합격자수 - 모의지원 상세보기에서 진학4에서 제시하는 추가합격자수를 의미합니다. 직접 세셔야합니다.

(6)표본 실지원자의 지원 정보- 모의지원 상세보기에서 표본이 어디에 지원할 지를 확인할 수 있는데 이를 수집하면 됩니다. 캡처보단 동영상 녹화하는게 낫습니다. 어차피 빵날 학과만 유독 볼거니 일일이 캡처할 필요도 없고, 다 보지도 않습니다.

우선은 이게 답니다. 글로하니까 적지만, 실제로 실제 대학의 학과는 20~30개가 기본입니다. 지원군이 많지 않은 서성중시아 나운데 나머진 정말 많습니다. 하지만 스나 절실히 붙고싶다면 하시는게 낫습니다. 업데이트 될 때마다, 아니면 최소 하루에 정기적인 시간대에 한번은 해야합니다. 만약에 추가적으로 해석할 유의할 데이터가 있으면 수집하십쇼 저건 예시입니다..

2.표본 수집후 해석한다

해석하는 지표는 스스로 찾아내면 좋겠지만, 제가 활용한 지표중 몇 개만 알려드립니다. 이런식으로 데이터를 해석하면됩니다. 인문논술 수리통계문제랑 비슷합니다.

다시 한번 말씀드리지만, 엑셀 함수 곱하기 나누기정돈 할줄 알아야 분석합니다.

●최초합격인원 차지비율=(4)진학4에서 제시하는 최초합격인원자수/총지원자수

- 진학4가 현재 합격인원중에서 최초합인원을 얼마나 잡았는지를 의미합니다.

=>가령 to가 10명인데 수집한 (4)의 자료에서 8명이라 잡았습니다. 그러면 8/10해서 0.8입니다. 아직 진학4는 2명이 더 들어올 것이라고 추정하는겁니다. 만약에 5명이라면, 5명 더 들어올거라고 생각하는것이고, 진학4생각엔 현재 지원풀이 적다고 판단하는겁니다. 반면 10명이나 11명이라면 지원풀이 다 들어왔을거라 예상하는 것이죠.

===>여기서 스나는 이런 가정에서 시작합니다. 값이 1미만이지만 실제로 이대로 간다면? 적정지원풀이 부족한게 아니라 이게 실제 지원풀이라면? 하지만 막상 해보면 많은 과들이 1미만이 뜬니다. 실제로 그래야하는게 표본추출이니 더 들와야하는게 맞습니다. 하지만 여기서 빵은 진짜로 1미만이 뜬 표본과 실제 원서지원자의 풀이 비슷하면 빵이날 가능성이 큼니다. 왜냐하면 가령 5명으로 잡은 근거는 대개는 누백을 사용한 적정 예측점수를 활용하는데, 실제 원서지원풀과 같다면 실제 최초합인원은 5명이아니라 10등인거죠. 진학4가 예측을 틀리게했다면 빵이터지는거죠.

==>이런 식으로, 계속 추이를 보면 여러 가지로 해석가능합니다. 계속 1 근처면 그 지원풀 그대로 원서까지 갈 가능성이 높겠죠.

=>보통은 gs성장이나 페it등 예측 점수랑 비교해서 풀의 적정성을 따지는게 바람직합니다. 다른 지표들도 마찬가지로...

●추합인원 차지비율=(5)진학4에서 제시하는 추가합격자수/(5)진학4에서 제시하는 추가합격자수+(4)진학4에서 제시하는 최초합격인원자수

- 진학4가 현재 추가합격자를 전체 현재 표본상 합격자수에 얼마나 잡았는지를 제시합니다.

=>가령 최초합인원(4)을 8명 잡았고, 추가합격자수(5)를 4명 잡았으면, 4/12해서 0.333이 나오겠죠. 일반적으로 다 통계내보면 0.333정도합니다.

==> 일반적인 다른 학과 평균보다 지나치게 높다면 현재 지원자수가 많이 빠져나갈 것이라 예측하는건데, 다소 주의해야할 게 현재 표본이 실제로 나갈지를 잘 봐야합니다.(3지망 지원정보를 봐야합니다) 실제로도 많이 빠져나갈 표본들이라면 추정이 맞겠지만, 그러지 않다면 과하게 커트를 낮게잡고 있는 것이죠. **빵이 나기 어려운 구조입니다.** 반대로 지나치게 평균이 낮으면 실제로도 덜 빠져나갈 것 같다면 맞는 추정이지만, 실제로 더 빠져나갈것같다면 이는 커트를 지나치게 높게 추정하고 있는 것이죠. **이런 경우 빵이 날 가능성이 있습니다.**

●표본상 합격비율=(5)진학4에서 제시하는 추가합격자수+(4)진학4에서 제시하는 최초합격인원자수/(2)실지원자수

-진학4에서 제시하는 총합격자수가 전체 지원자수 대비 얼마나 되는가를 의미합니다.

=>가령 최초합8명+추가합격자4명이고 전체 실지원자수가 36명이면, 12/36해서 0.333입니다 일반적으로 0.35정도 잡히는 것 같습니다.

==>이제 지나치게 작은 수치라면, 최초합격인원차지비율처럼 더 들어올 수 있는데 지나치게 커트를 높게 측정한걸 수도 있습니다. 하지만 적정하게 측정될 수도 있는데 최초합격인원차지비율이 1에 가까운데 추합인원이 낮은 경우엔 표본상합격비율이 낮을 순 있습니다. 복합적으로 고려를 해야합니다.

==>**빵시나리오는 최초합격자비율도 낮는데 표본상합격비율은 높게(혹은 평균과 비슷한) 잡힌 경우입니다.** 이 경우는 집단이 다 들어오지 않았다고 판단하지만 그럼에도 다른 학과와 비슷하게 표본합격비율을 잡았더라면, 실제 합격 커트는 더 내려갈 수 있는 겁니다. 최초합인원이 더욱 넉넉해질 수 있는거니까요.

●지원인원 추정=(3)진학4에서 제시하는 각과의 예상경쟁률*모집인원(TO)

-진학4가 추정하는 예상경쟁률에따라서 몇 명이 지원할까를 의미합니다

=>예를들어 TO가 10명인데 예상경쟁률을 6이라 잡아봤으면, 지원인원추정은 60명입니다

=>일반적으로, 실지원자수에 몇 퍼센트 더 붙어서 지원을 합니다. 그런데 이 지원풀에 미달하는 점수로 0.5배수~1배수만큼 더 지원을 합니다. KY에 갈수록 1배수에 가깝고 인기학과의 경우엔 더 그렇습니다. 소위 불나방이죠. 안될 걸 알지만 기반아가는...아니면 17연심리를 노리는.. 그런 경우죠

==>의미는 이렇습니다. 지원인원추정은 60명이고, 현재 실지원자수는 30명입니다. 만약 이과가 KY에 해당한다면 1배수의 허수(실지원자수보다 못한 점수)들이 지원합니다. 그럼 실제로 이 풀이 그대로 지원할 갈 가능성이 높고, 실제 지원풀도 이대로갈 가능성이 높습니다.

==> 이제 스나 시나리오는 이겁니다. **지원인원추정대비 실지원자수는 적정한것같은데(지원인원추정대비/실지원자수가 0.5~0.6이면 적정하다고 생각하면됩니다), 다른 지표들은 여전히 숫자가 부족하다고 추정하는 경우입니다.** 왜냐면, 지원자수는 적정한데 실제로 더 많이 들어올 거라고 기대하는 가정 자체가 틀릴 수 있기 때문입니다.

====>강력한 힌트지만, 지금보니 놀랍게도 이렇게해서 작년 빵 몇 개가 설명이됩니다. 잘 가져가셨으면 좋겠습니다. 물론 몇 개는 설명이 안됩니다. 이런 것은 경쟁률이나, 2편에서 설명한 여러 요소들이 복합적으로 작용해서겠죠.

====>예상경쟁률을 기초로해서 진학4가 많이 값들을 도출하는 걸로 알고있습니다. 예상경쟁률을 믿어보되, 여러 데이터간 모순을 찾아보는겁니다.

1.그 외 여러지표를 충분히 만들어 볼 수 있습니다. 이걸 대표적인 지표고, 저같은 경우 예상 등수까지 추론했는데 적어도 스나로 지원한 과에선 맞췄네요.(이건 IF함수랑 여러 가지 가정을 더 추가하고 추정치 변동을 가미한거라 글로설명하기 너무 어렵네요) 당시에 그렇게 절박했나 봅니다.

2.또 KY나 S는 당연히 헤드카운팅할수록 더 정확해질겁니다. S스나는 잘 모르겠으나 KY스나의 경우 표본 전부기록(점수대까지)하고 얼마나 변동이 이뤄지는지 잘 보는게 낫습니다. 하지만 그럼에도 스나 찾기는 너무 어렵겠쥬.

====>KY의 경우엔 솔직히 장난치기 너무 쉬운데 잡아내기 너무 어렵습니다. 당장 수시합격자 5명 계정만 가지고가서 허위점수로 허위점공하면 됩니다. 아무리 CON선텩이니 표본분석이니 할 것이 없는게, 직접 표본을 조작하면되니까요. 15연자전이든, 16S경영 미수사건이든, 시장을 조작하는건 꽤나 유례가 깊습니다. 지금까지 KY 핵빵들도... 그런 것들이 흑여 있지 않을까 조심스레 생각해봅니다. 특히나 지금같이 특정 플랫폼에 지나치게 기대고 있는 상황에선 더더욱 효과가 크쥬. 상위권에서 인터넷 안되는거아니면 누가 진학4안끊고하나요? 하지만 저는 여전히 잡아내는 법을 모릅니다. 누군가는 잘 찾아내서 그런 '세력'을 잘 찾아내서 잘 들어갔음 좋겠습니다. 아무래도 KY미만부터는 이런 조작이 값어치가 없다판단하는진 잘 모르겠습니다만 이런 핵빵이 잘 안나는 것 같습니다.

3.분석의 핵심은 어느정도 커트라인까진 피고 있어야합니다. 이런식으로 해석하고, 추합과 불합의 경계선을 진학4에선 XXX점이라 추정하고 GS성장은 YYY점이라 추정하고있으면 어느게 더 맞을지, 그리고 실제 만든 지표상과 비교로 해봤을 때 실제론 몇 점이 될지 표본정보(다른 표본이 어디로 빠질 건지와 같은 정보와 함께) 추정을 해야합니다. 이런식으로 상위대학부터 피고 자기가 생각하는 미니멈대학까지 점수대 추정이 얼추 되고, 현재 표본상 추합인원까지 고려를 하면 이제 이과가 스나가 될 학과인지 아닌지(적어도 스나로 써볼만한 확률이 나올지) 나옵니다.

4.이 분석은 기본적으로 횡적으로 같은 날 다른과와 비교해봐야하고, 종적으로 매일 추세를 읽어야합니다. 하다보면, 실지원자수가 부족하니 더 들어가는 과가 있고, 반대로 많다고 생각해서 빠지는 과가 있고... 정시이월인원이 늘어나서 변동이 생길 수 있고.. 계속 이런걸 지켜 봐야합니다. 물론 실제 지원 풀은 최종날에 가까울수록 표본과 가까워 집니다. 그럼에도, 경쟁률이나, 작성한 표본추이들, 그리고 2편에서의 여러 요인들이 중첩해서 진학4 최종표본과 다르게 실지원이될 겁니다. 이걸 최종적으로 예측하면 되는겁니다. 경쟁률만이 유일하게 논의하지않은 요인이므로 이를 4편에서 다루겠습니다.

<경쟁률>

일반적으로, 최종경쟁률이 낮을수록 합격 커트는 일반적으로 내려갑니다. 당연하게도, 지원자 풀이 적으니까요.

=>하지만 실제로 깐다면, 다릅니다. 3편에서 표본의 허수와 실지원자수 비율이 1:1정도된다고 했는데 같은 2:1의 경쟁률인 두 과가 있다고 합시다. 그런데 한 과가 실지원자로만 모여있게 지원이된다면 실제 경쟁률은 한 쪽은 1(허수제외):1이고, 다른 한쪽은 2:1이겠쥬? 반대로 10:1이라 하더라도 실지원자수대 허수 비율이 1:9라면, 실질적인 경쟁률은 1:1일겁니다. 작년에

20연경제와 20연아과 경쟁률이 실제로 그렇습니다. 20연경제는 2:1에 가까웠고, 20연아과의 경쟁률은 10:1이었지만, 실제로 커트에는 큰 이상이 없었습니다.

더구나 최종경쟁률은 원서 쓰고나서 아는 경쟁률이므로, 원서쓰고나서는 이제 스나 관점에서 사실상 의미가 없습니다. 쓰고나서 주변에서 얼마나 쫓는지는 쓸 때 고려하는게 아니죠. 실시간 경쟁률을 통해서 알고 싶은건, 경쟁률을 통해서 과연 얼마나 실지원자들이 몰렸는가가 중요합니다.

보통 두 가지 지표로 실지원자가 몰렸을 것이라고 예측 가능합니다.

1.마감일 전날까지의 경쟁률-아주 일반적으로, 마감일 전날까지의 지원은 안정을 깔고 쓰는 경우가 많습니다. 경영이나 경제가 특히 초반 경쟁률이 높습니다.

==>유익한 해석이 나오려면 과거 경쟁률 추이와 비교해야겠지요. 만약에 같은 학과인데 20년도엔 마감일 전까지 0.5:1의 지원을 보이다가 21년엔 1:1을 보인다 이러면 폭의 조짐이 보입니다. 하지만 반대로 빵의 시그널까지 나옵니다. 왜냐면 최초합하여 추합하는 인원이 그만큼 늘어난 걸수도 있는데, 도리어 일반 추합인원들이나 스나, 허수들이 높은 경쟁률을 보고 오히려 더 쪼는 시나리오가 날 수 있습니다.

====>이런 판별은 보통 표본에서 나타납니다. 표본이 과하게 몰릴 것 같다는 분석을 최종적으로 했는데 실제 1:1이 된다면 그게 현실이된 것이고, 과하게 몰리지 않을 것이라 봤는데 갑자기 들어온 것이라면 숨어있던 표본들이 여기저기서 몰려왔다고 해석이 가능합니다. 하지만 어느 경우건 마감일 전날까지의 높은 경쟁률은 일반 추합인원들이나 스나, 허수들에겐 호재가 아니라서 쫓게 하는 유인이기도 합니다.

====>하지만 이런걸 노리고 소수과나 일반적으로 선호도가 낮은 경우는 허수가 초장부터 들어가기도 합니다. 스스로를 뺑카로 세우는 거죠.

2.마감일의 마지막경쟁률-보통 14:00나 15:00에 직전경쟁률을 발표합니다. 발표 이후 이 때 기지원자만큼 지원을 보통을 합니다. 가령 1.5:1로 14:00에 직전경쟁률이 발표가되면, 보통은 최종경쟁률은 3:1입니다.

====>이것도 마찬가지로 허수와 실지원자수의 비율이 1:1이다란 가정과 호응합니다.

====>직전경쟁률이 타과보다 적으면 엄청 몰리고, 타과보다 많으면 덜 몰리는 성향이 있습니다.

====>어쨌건 스나의 관점에서는, 너무 극단은 피하는게 좋습니다. 너무 덜 몰린 곳엔 여차피 스나든 추합인원이든 허수든 다 몰리고, 너무 많이 몰린 곳엔 마찬가지로 이미 최초합인원과 추합인원이 많이 모였다고 판단할 여지가 있습니다. 하지만 정말로 쓸 곳이없다면 이런데라도 써야겠죠.

====>마지막까지 유지해야할 태도는 '뺑구인 곳을 내가 씩니다.' 표본 분석상 뺑구일 학과는 실제 직전경쟁률까지의 추이에서도 분석이 틀리지 않은걸 대체로 보여줍니다.(추이는 뺑구 시나리오마다 다릅니다...) 그런데 분석을 어긋나게하는건 바로 직전경쟁률의 자료입니다. 표본 분석 최종 자료는 마지막날에 업데이트가 안되고 그 전에 업데이트로 마감하여 모두가 들고있는 자료가 같습니다.

또 직전경쟁률도 뺑구조짐이라고 나만 생각하면 괜찮은데, 스나들도(그리고 추합인원들도) 그

렇게 생각하고 모두가 써버린다면 그 과는 정상 입결이나 폭으로 갑니다.

실제로 저의 경우 붙은 대학에 3개의 과가 직전경쟁률까지보면 빵이날 것이라 봤는데, 한 과는 가도 너무 고생할 과일거 같아 거르고, 나머지 두과에서 고민했는데, 결국 너무 빵인 것 같은 과를 거르고 나머지 과를 택했습니다. 결국엔 고생할 과는 핵빵이 났고 너무 빵인 과는 그 대학 입결1위를 찍었습니다.

=>항상 유념하세요. 스나는 두 과정입니다. 빵꾸인 곳을 찾고, 빵꾸인 곳에 붙는 두 과정입니다. 모두가 빵꾸를 찾아내고 모두가 지원한다면 내가 붙기 어려워집니다. 빵꾸는 원서마감때 결정되지, 직전경쟁률이 발표된 시점에 결정되는게 아닙니다.

===>2시간~3시간 이내로 최종결정해야하는데, 저는 쉽지 않았습니다. 항상 정시원서 쓸 때 다군 제외하곤 전부 마감직전 5분이내... 짧은건 1분이내 했던 것도 기억납니다. 빠른 해석이 필요합니다. 왜냐면 스나는 특정 대학에 걸쳐있는게 아니라 특정 라인에 걸쳐서 하기때문이기 에, 경쟁률을 죄다 분석해야합니다.

특히나, 올해 같이 응시인원감소의 이슈가 있는 상황에서 상위대학에서 직전경쟁률이 작년에 비해 여유롭다면, 추합여력이 더생길 수 있다고 판단가능합니다. 그래서 결국 표본 추이든, 경쟁률 추이든 S부터 본인 원하는 대학라인까지 다 봐야한다는 겁니다. 특히 본인이 쓸 스나학과가 추합인원이 많이 생기는 과라면 더더욱이겠죠. 이런 분석까지 다 해내야 확실히 어느정도 보입니다.

<마무리>

1.그냥 이대로 머릿속에만 담아두긴 좀 뭐해서 시작했던 것인데... 역시나 쉽지 않네요. 얼마나 도움이 됐을란진 잘 모르겠습니다. 논쟁의 여지도 있는 부분도 많고, 전달이 잘 매끄럽지 않은 부분도 있을 것 같구요. 그렇다고 돈 받고 하는 것도 아니니 막 엄청 신경써서 작성하기도 뭐 하구요. 좀 빨리 써서 포스팅하는 건데 차일피일 미루다가... 작성이 많이 늦게된 것도 유감입니다. 자랑하려고 올린 글도 전혀 아니고 제 입시과정은 많이 부끄럽기도 합니다. 부끄러운 과정을 괜히 보여주나 싶기도 하고 많은 사람들은 그저 스크롤을 내리겠지만 그럼에도 분명히 누군가에게, 그리고 이 내용중 아주 조그마한 부분이라도 도움이 되지 않을까...라는 기대에 올려봅니다.

2.a4 22쪽 분량이 나왔는데, 제가 글을 하도 길게쓰고 장황하게 쓰는지라 핵심만 잘 가져갔기를 바랍니다. 게다가 맞춤법 검사도 하지도 않은 지라 대충 한국인이라면 잘 알아드으셨길 바랍니다..

3.스나하면서 데이터 찾고 분석하는 과정들이 힘들겁니다. 기분 좋게 올 만점이거나 한 두 개 틀렸으면 기분 좋게 원서를 쓰셨겠지만, 그러지 못했으니 고생해야죠... 특히 스나라는 것 자체가 자기 점수를 일단은 부정하고 여러 예측을 부정하는 근거를 찾는 것이니 더 힘들고, 더 더욱 힘든 것은 잘 본 사람들의 성적표를 보거나, 또 여러 온라인이든 오프라인에서의 수시합격자들을 더욱 힘들겁니다. 어쩔 수 없어요. 참고해야죠...

4.결국 수능은 수단입니다. 정시전형에서 '대학합격'이라는 목적을 위한 수단이지 그 자체가 목적이 아닙니다. 평가원에서든 교육부에서든 왜 그냥 성적순대로 합격시키지 않을까란 질문엔, 결국 이런 스나의 기회를 열어주는 측면도 있다 생각합니다. 마지막 패자부활전(패자라는 말도 참 민감합니다만)인 것이죠. 어찌되었든 마지막 기회입니다.

5.그리고 살아보니, 대학도 결국 인생의 목적이아니라 수단입니다. 대학 잘간다고 행복한 것도 아니고 못간다고 불행한 것도 아니더라구요. 대학이 전부일 시기이지만, 너무 못봐서 스나하는 처지 너무 비판하지마세요. 인생은 여러 기회가 있으니까요. 건투를 빕니다.

6.무단배포를 하든 알아서들...하시길 이걸 가지고 뭘 취업한다든가 현금화할 생각 전혀 없습니다. 제가 돈 받은 적도 없으니 아주 흑시나 저한테 의존하지마시고... 스스로를 믿으시길.