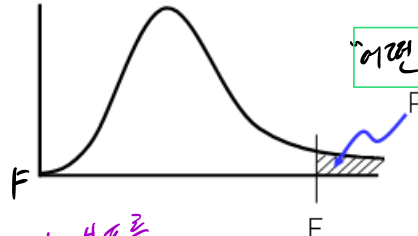
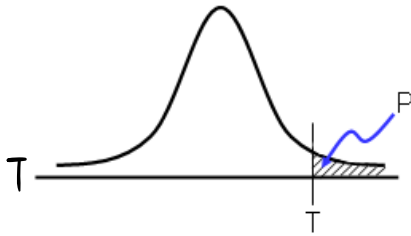


StatEdu *207.22.67

<http://www.statedu.com/736/>

✓ T-value와 F-value가 본포를 따는 이유:
 ✓ 포본이 달라짐에 따라, T-value와 F-value도 달라진다.



“어떤 사건이 우연히 발생할 확률”

통계분석을 하다보면 T, F, χ^2 등의 값이 자주 나오게 되는데, 항상 P 값과 같이 나오게 된다. 이때, 이 값들이 무엇을 뜻하는지는 paper 에도 나오지 않고, 오직 P 값만으로 설명을 하게 된다.

T-test 를 하게 되면 T, P 값이
 ANOVA 를 하게 되면 F, P 값이
 회귀분석에서는 F, T, P 값이
 교차분석에서는 χ^2 , P 값이 나오게 된다.

다 분포를
 비교한다!!!
 = P(유의 확률)은
 이래서든 존재함.

T, F, χ^2 값이 나와야 P 값이 나온다!

T 값에 대한 확률 or
 F 값에 대한 확률

위의 그림을 보게 되면 x 축에 T, F 값이 있다. 그리고, 그 값을 기준으로 오른쪽의 면적이 P 이다.

컴퓨터가 발전되기 전까지는 통계분석을 할 때, 손으로 일일이 계산하였다. T-test 의 경우에는 2 그룹간의 차이가 있는가를 알기 위해서 T 값을 계산하였다. 그렇게 계산되어진 값이 검정통계량이라고 하며, 통계분석에서 출력되는 T 값이다. 이 T 값을 가지고 어떤 기준치(기각역, 임계치)라고 할)와 서로 비교를 하게 되는데, 그 기준치는 유의수준이 0.05가 되게 하는 값이다. 즉, 위의 그림에서 오른쪽의 빗금친 면적이 0.05이게 하는 x 축의 값이다.

이렇게 계산되어진 T 값이 기준값보다 크면 H1 을 선택해서 유의한 차이가 있다라고 하며, 그 기준값보다 작으면 유의한 차이가 없다라고 하는 것이다.

대립가설

H0(귀무가설)을
 선택하여

그런데, 이 기준값은 분석 할때마다 변하게 된다. 정확하게는 case 의 수 등에 따라 다르다는 것이다. 그래서, 분석을 하게 되면 매번 그 기준값을 찾아주어야 하는데, 그 기준값을 정리해 놓은 것이 바로 통계학 책의 맨 마지막에 있는 수표들이다. 결국, 손으로 계산한 T 값과 수표에서 찾은 기준값을 서로 비교해서 유의한 차이가 있는가 없는가를 결정하는 것이다.

그러다가, 컴퓨터가 발전이 되면서 직접 어떤 T 값이 계산되어 나왔을 때, 그때의 빗금친 면적을 컴퓨터가 계산하게 되었다. 이렇게 계산되어진 값이 바로 P 값이다. 이 P 값 역시 어떤 기준치와 비교를 하여야 하는데, 그 기준치가 앞에서 나왔던 0.05인 것이다.

이렇게 P 값을 계산하게 되면 기준치가 변하지 않고 항상 같은 0.05란 값과 비교를 하게 되므로 결과의 해석을 하는데, 상당히 편하게 되었다.

X. 유의확률 = $p\text{-value}$

· 일반적인 유의수준 : 0.05 (5%)

· 기각역 : 귀무가설을 기각해야하는 검정통계량의 구간

· 유의미한 차이: 유의확률이 유의수준보다 작으면,
'평균1'과 '평균2'의 차이가
유의하다고 한다.

· 통계적으로 유의한 차이가 있다고 해서, 실제로 의미있는
차이라고 하는 것이 아니다. 통계적으로 '0.001'
만큼 유의한 차이가 있다는 결과가 나왔더라도,
실제 상황에서는 이 정도 차이를 무시해도 상관
없을 수 있다. 그러면, 이것은 '무의미한 차이'인
것이다.