





아웃라이어 상위 5개

[22222,13222,21222,22111,31111]

아웃라이어 하위 5개

[23211,13211,22322,11211,32122]

두 path에 발생한 총 불량률 수의 평균 비교



평균이 다르다고 할 수 없다.

[19,166]

[1,50]

Chi square 검정

p value:0.111

outlier상위 5개와 하위 5개 중 200개를 넘는 불량률에 대해

분산이 다르다고 할 수 없다.

하지만~~~~

20개 이상에 데이터에 대해서

상위 다섯개의 path

[22222,13222,21222,22111,31111]

하위 다섯개의 path

[11311,23311,22333,13333,32311]



두 값이 정규성을 따르지 않으므로

mans whitney method에 의해

평균은 다르다고 할 수 없다.

[19,166]

[4,126]

Chi square 검정

p value:0.028<0.05

20이상 데이터에 대해 상위 5개와 하위 5개의 분산은 다르다고 할 수 있다.

20개 미만에 데이터에 대해서



p value>0.05이상이므로 불량갯수의 평균차이가 있다고 할 수 없다

[15,117]

[3,108]

Chi square 검정

p value:0.020

20미 데이터에 대해 상위 5개와 하위 5개의 분산은 다르다고 할 수 있다.