

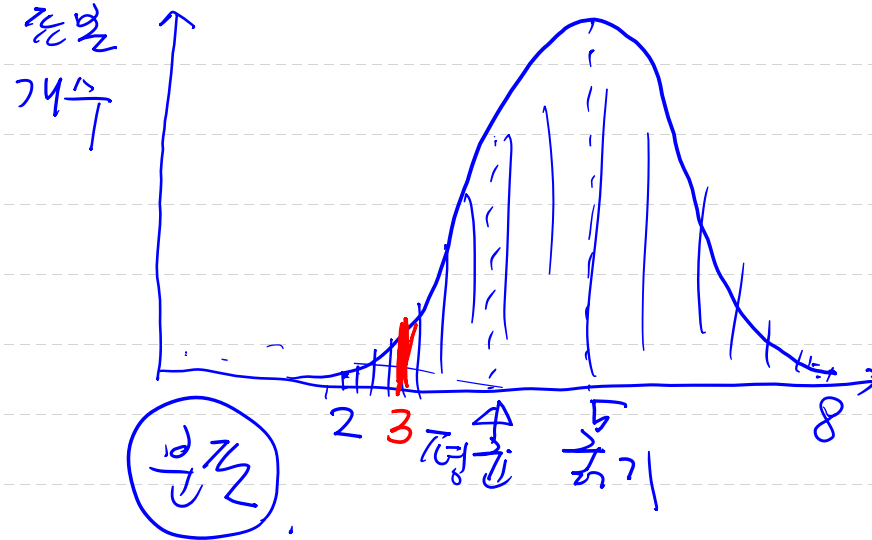
모집단, 표본집단. (샘플)
 평균, 분산.

P-value.

모집단 분포 (모집단), 한 점 모집 (표본)

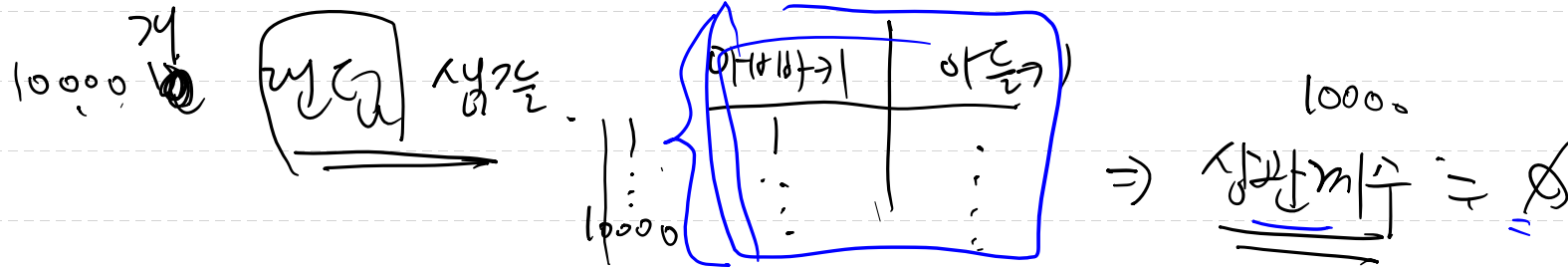
• 표본 → 모집 한 점 → 표본 : 3

\Rightarrow (1000000) 한 점 모집 분포를 10000000 가려냄.

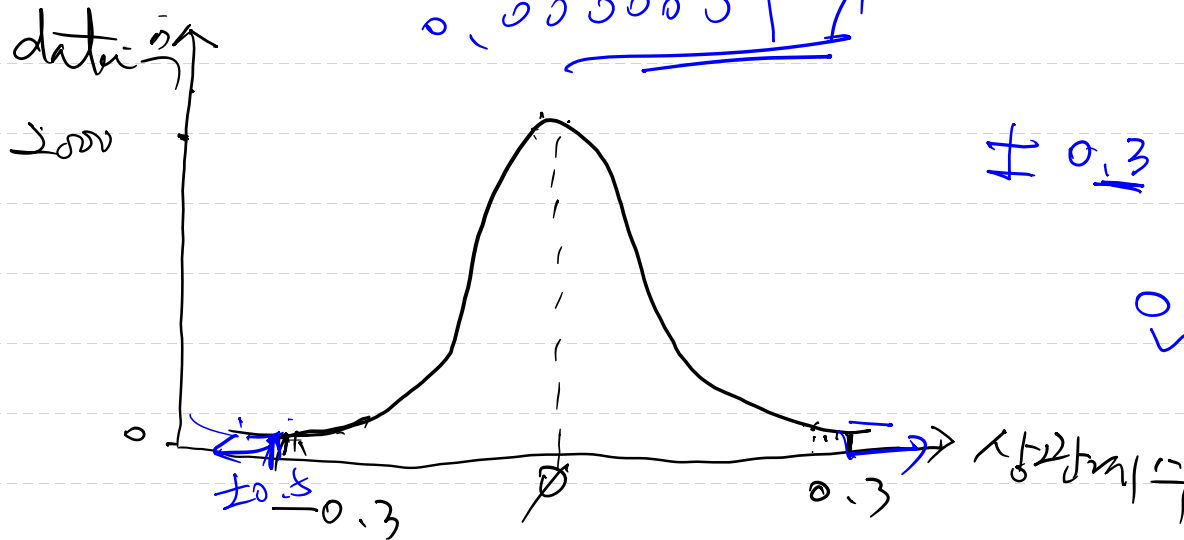


아버지키와 아들키의 상관계수 (연속형 변수)

- 상관계수 : 0.5 의미 있는 값? \Rightarrow "정확 이상"



1000000 상관계수의 \Rightarrow 0.000001



$\pm 0.3 \uparrow$ 상관계수가 22%

0.22% 확률

AB 테스트 (비구조적 변수)

	메뉴 A	메뉴 B	총계
반응	35	25	60
부시	15	25	40
총계	50	50	100

"메뉴 A가 효과가 있다"

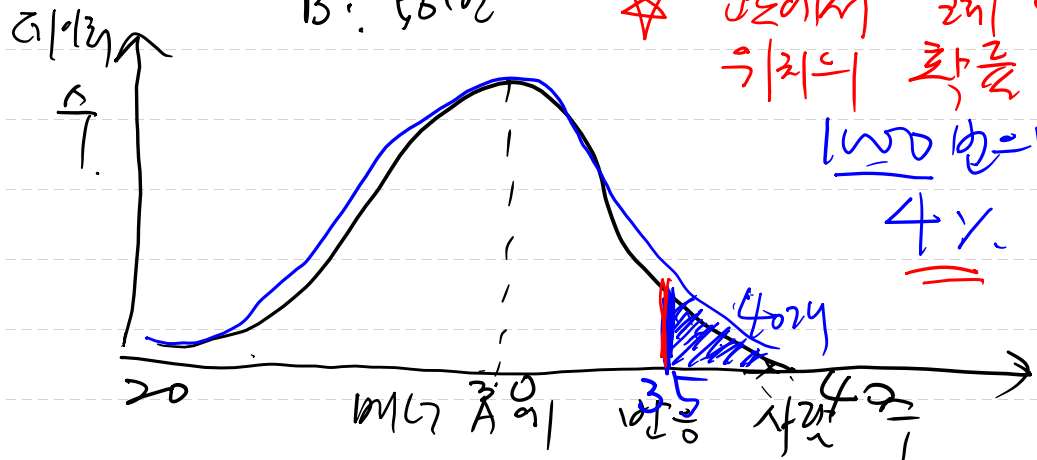
부정 "우연이 아니다"

영향력 2 이상 10 이상

결과	메뉴 종류	반응
A	반응	
A	부시	
B		
...		

A: 50회
B: 50회

	A	B	총계
반응	a	60-a	60
부시	50-a	a-10	40
총계	50	50	100



분포에서 실제 데이터의 상대적인 위치의 확률 값: p-value.

100번의 실험에서 35보다 큰 경우가 4%라면, 40개 실험에서 35보다 큰 값이 나올 것이다.

정규분포 p-value?

0.3 $\pm 0.3 \uparrow$ 10000개 22건. p-value 0.0022.

@ 0.22%: 매우 드물다

- P-value: 데이터 중 차이가 있거나. 제한지 or 제한하지의
경로들 $\phi \sim 1$ 사이 어느 정도.

< 1: 차이가 없는 것.

< ϕ : 차이가 너무 적음. 차이를 인정하는 수준.

- 5% 기준: 유의 수준. (데이터 간 차이가 제한하지 정도)

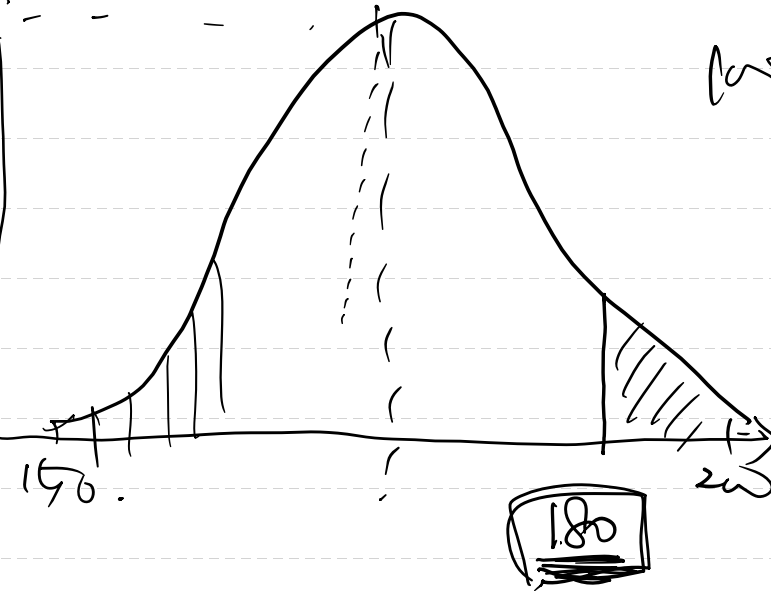
- 유의성 검정. $p\text{-value} < \text{유의수준} \Rightarrow$ 통계적으로 유의미함.

$p\text{-value} > \text{유의수준} \Rightarrow$ 우연치대로 충분히 비슷할
수 있는 차이.

아빠가, 아들기.

아들기

0.05



1000명 중, 200명

\Rightarrow 0.2

아들기

가우스 분포 (정규 분포)

◦ 정규 분포 확률 밀도 함수 이라는. \hookrightarrow 평균, 표준편차에 따라.

분포 함수의 모양이 결정된.

◦ 문제 : 분포의 모양을 이름 결정하는 값

◦ 모집단? 통계적 관찰 or 처리하는 대상 집합.

◦ 연구조사? 모집단에 대해 연구 조사

ex) 여론조사, 자연. 개체군 연구.

◦ 통계 추출? 집단의 특성을 추출 (평균, 분산 등)
(Sampling)

통계 (Sample) 추출 → 조사
↑ 통계적 추출.

◦ 통계조사? 통계를 조사하는 작업.

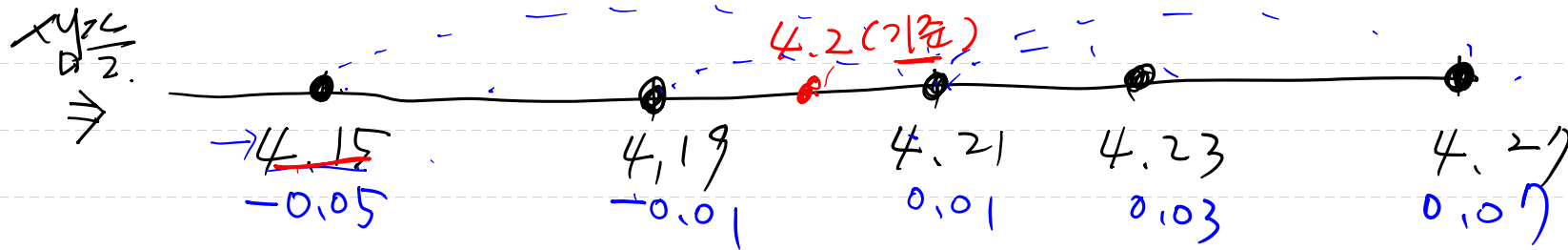
통계평균: 추출된 통계를 산술 평균, $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$

통계분산: $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

⇒ 통계평균, 통계분산.

$\bar{x} = \frac{(-0.05) + (-0.01) + 0.01 + 0.03 + 0.07}{5} = \frac{0.05}{5} = 0.01\%$
 $s_x^2 = \frac{(-0.05 - 0.01)^2 + (-0.01 - 0.01)^2 + \dots + (0.07 - 0.01)^2}{5}$

2. 평균을 도우 4.2% 편미와 수 있음. $\frac{0.008}{4} = 0.002$



5개 data의 평균을 도우 4.2% 볼 수 있음가?

\Rightarrow 기준 편미

기준 편미 : 편미 편미 ϕ 인리 아인리 편미.

① 각 data - 기준

② 편미 편미, 편미 편미, 편미 편미 편미 편미. $s = \sqrt{\frac{0.002}{5}} = 0.02\%$
 $\frac{4.15 + 4.19 + \dots + 4.27}{5} = 4.2$

$$\frac{t_{\text{값}}}{(t_{\text{값}}/5)} = \frac{0.01\%}{0.02\%} = 0.5$$

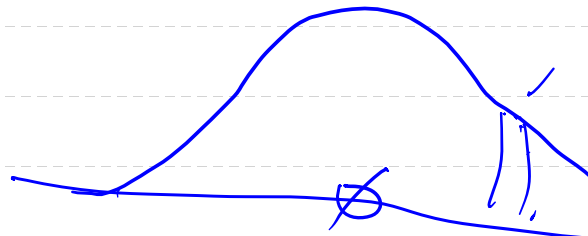
$\begin{cases} \bar{x} = 0.01, \\ S^2 x = 0.02\% \end{cases} \Rightarrow$ 평균 값을 4.2% 보다 0.01 크다

\Rightarrow $t_{\text{값}}$

집중화
 의미는 동일

기준값으로부터 상대적 수준에 따라 있는가?

$t_{\text{값}}$



반대로 - 평균
 표준편차 - 기준값
 표준편차의 표준편차
 표준편차

