

Data Analysis

저온숙성 사전과 사후 딸기 당도 비교

대응된 두 집단 모평균 비교
대응표본 t검정

학습효과

- 저온숙성 전후의 당도를 비교해서, 당도가 향상되었는지 검정할 수 있다.

표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09

1. 확률변수 사전 당도와 사후 당도로부터 사후 당도 - 사전 당도 차이라는 새로운 확률변수를 이끌어 낼 수 있다.
2. 새로운 확률변수로부터 검정통계량을 계산해서 구할 수 있다.
3. 검정통계량에 적합한 확률분포로부터 유의확률을 구할 수 있다.
4. 유의확률과 유의수준을 비교해서, 가설을 검정하고 결론을 정리할 수 있다.

학습순서

1. 새로운 확률변수 : 대응된 두 확률변수 차이
2. 가설검정 : 대응된 두 집단 모평균 비교 - 대응표본 t검정

확률변수 : 사전 당도, 사후 당도

딸기ID	사전 당도(Brix)	사후 당도(Brix)
1	12.8	13.4
2	12.0	11.9
3	12.1	12.4
4	13.3	13.3
5	13.2	13.7
6	12.7	12.8
7	11.9	12.0
8	12.1	12.3
9	12.2	12.3
10	11.8	12.0
11	12.6	12.9
12	12.9	13.0
13	12.1	12.7
14	12.0	11.9
15	12.7	13.2
16	11.6	11.6
17	12.1	12.4
18	11.8	11.6
19	12.5	12.8
20	12.1	12.2

새로운 확률변수 : 사후 당도 - 사전 당도 차이

딸기ID	사전 당도(Brix)	사후 당도(Brix)	사후 당도 - 사전 당도 차이
1	12.8	13.4	0.6
2	12.0	11.9	-0.1
3	12.1	12.4	0.3
4	13.3	13.3	0.0
5	13.2	13.7	0.5
6	12.7	12.8	0.1
7	11.9	12.0	0.1
8	12.1	12.3	0.2
9	12.2	12.3	0.1
10	11.8	12.0	0.2
11	12.6	12.9	0.3
12	12.9	13.0	0.1
13	12.1	12.7	0.6
14	12.0	11.9	-0.1
15	12.7	13.2	0.5
16	11.6	11.6	0.0
17	12.1	12.4	0.3
18	11.8	11.6	-0.2
19	12.5	12.8	0.3
20	12.1	12.2	0.1

새로운 확률변수 : 표본평균

딸기ID	사전 당도(Brix)	사후 당도(Brix)
1	12.8	13.4
2	12.0	11.9
3	12.1	12.4
4	13.3	13.3
5	13.2	13.7
6	12.7	12.8
7	11.9	12.0
8	12.1	12.3
9	12.2	12.3
10	11.8	12.0
11	12.6	12.9
12	12.9	13.0
13	12.1	12.7
14	12.0	11.9
15	12.7	13.2
16	11.6	11.6
17	12.1	12.4
18	11.8	11.6
19	12.5	12.8
20	12.1	12.2

사후 당도 - 사전 당도 차이
0.6
-0.1
0.3
0.0
0.5
0.1
0.1
0.2
0.1
0.2
0.3
0.1
0.6
-0.1
0.5
0.0
0.3
-0.2
0.3
0.1

사후 당도 - 사전 당도 차이 평균	확률분포
0.20	정규분포 가정

검정통계량

표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09
n	\bar{X}_1	\bar{X}_2	\bar{D}	df	S_D^2	S_D	$SE(\bar{D})$	t	p	α	CV

검정통계량

표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09
n	\bar{X}_1	\bar{X}_2	\bar{D}	df	S_D^2	S_D	$SE(\bar{D})$	t	p	α	CV

$$D = X_2 - X_1$$

$$\bar{D} = \bar{X}_2 - \bar{X}_1$$

OR

$$\bar{d}_i = \frac{\sum_{i=1}^n d_i}{n}$$

검정통계량

표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09
n	\bar{X}_1	\bar{X}_2	\bar{D}	df	S_D^2	S_D	$SE(\bar{D})$	t	p	α	CV

$$S_D^2 = \frac{\sum_{i=1}^n (d_i - \bar{d}_i)^2}{n - 1}$$

검정통계량

표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09
n	\bar{X}_1	\bar{X}_2	\bar{D}	df	S_D^2	S_D	$SE(\bar{D})$	t	p	α	CV

$$SE(\bar{D}) = \sqrt{\text{Var}(\bar{D})} = \sigma_{\bar{D}} = \sqrt{\frac{\sigma_D^2}{n}} \sim \sqrt{\frac{S_D^2}{n}} = \frac{S_D}{\sqrt{n}}$$

검정통계량

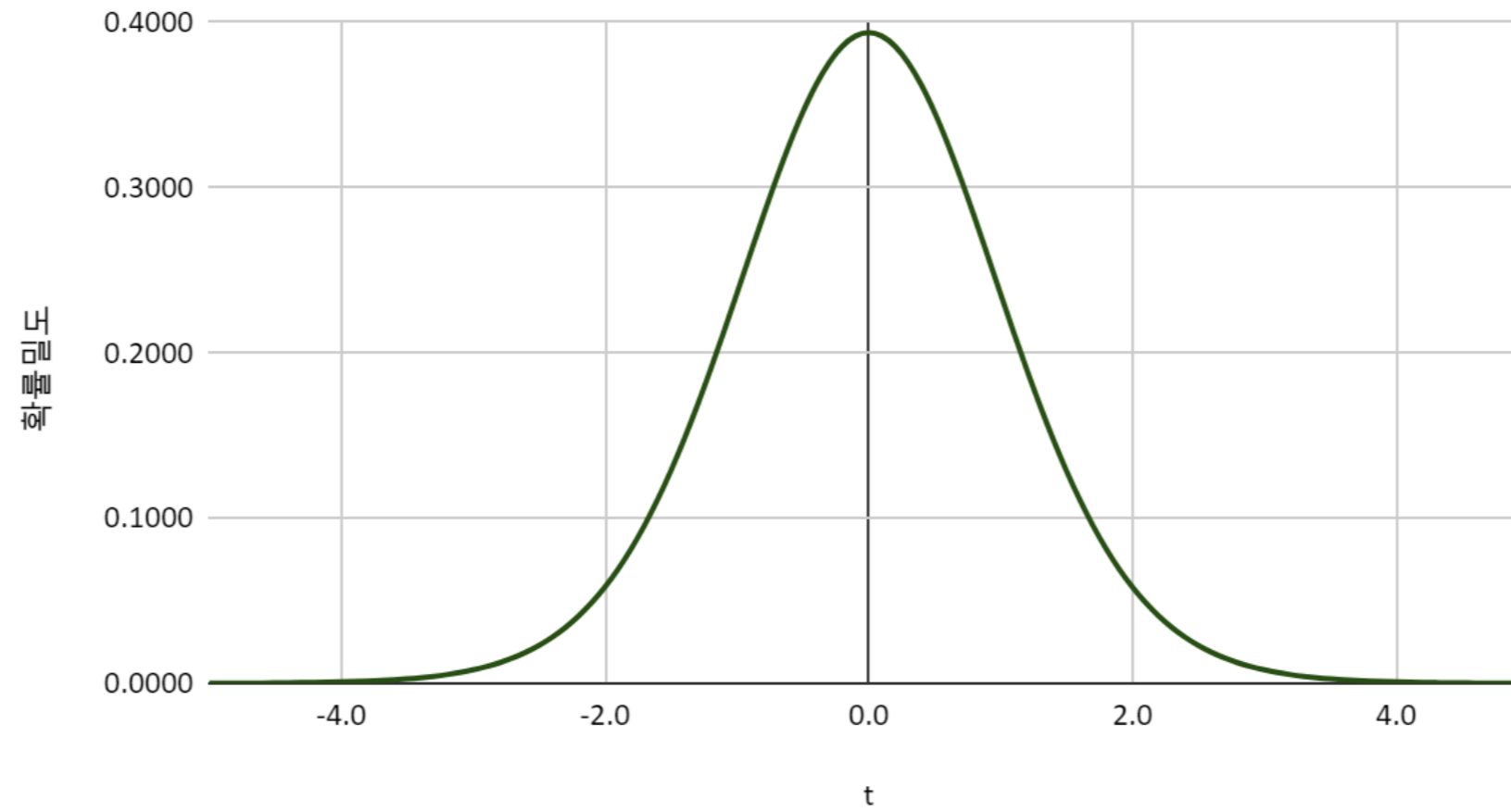
표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09
n	\bar{X}_1	\bar{X}_2	\bar{D}	df	S_D^2	S_D	$SE(\bar{D})$	t	p	α	CV

$$t = \frac{(\bar{X}_2 - \bar{X}_1) - D_0}{\frac{S_D}{\sqrt{n}}} = \frac{\bar{D} - D_0}{\frac{S_D}{\sqrt{n}}}$$

확률분포

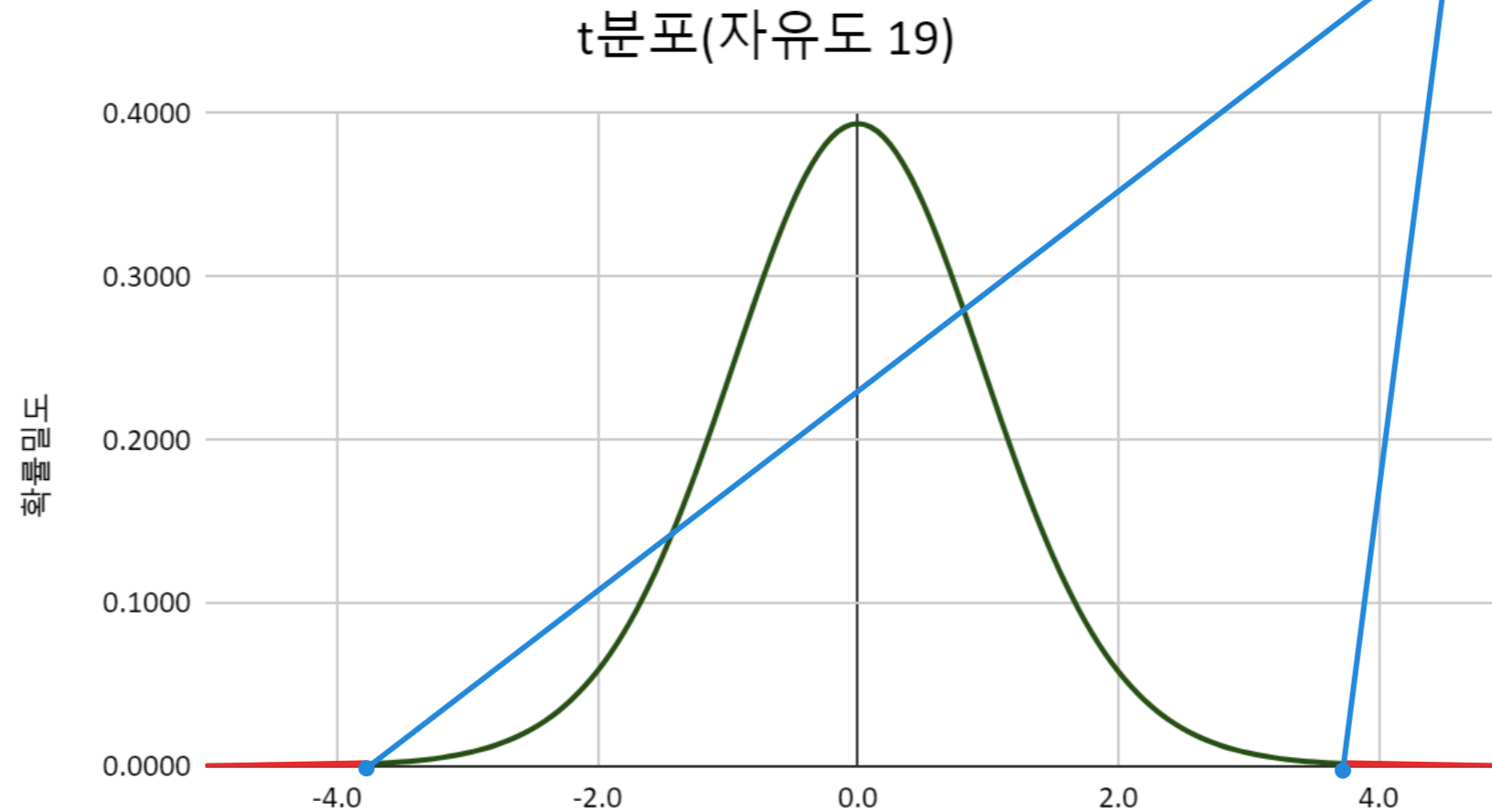
표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09

t분포(자유도 19)



유의확률

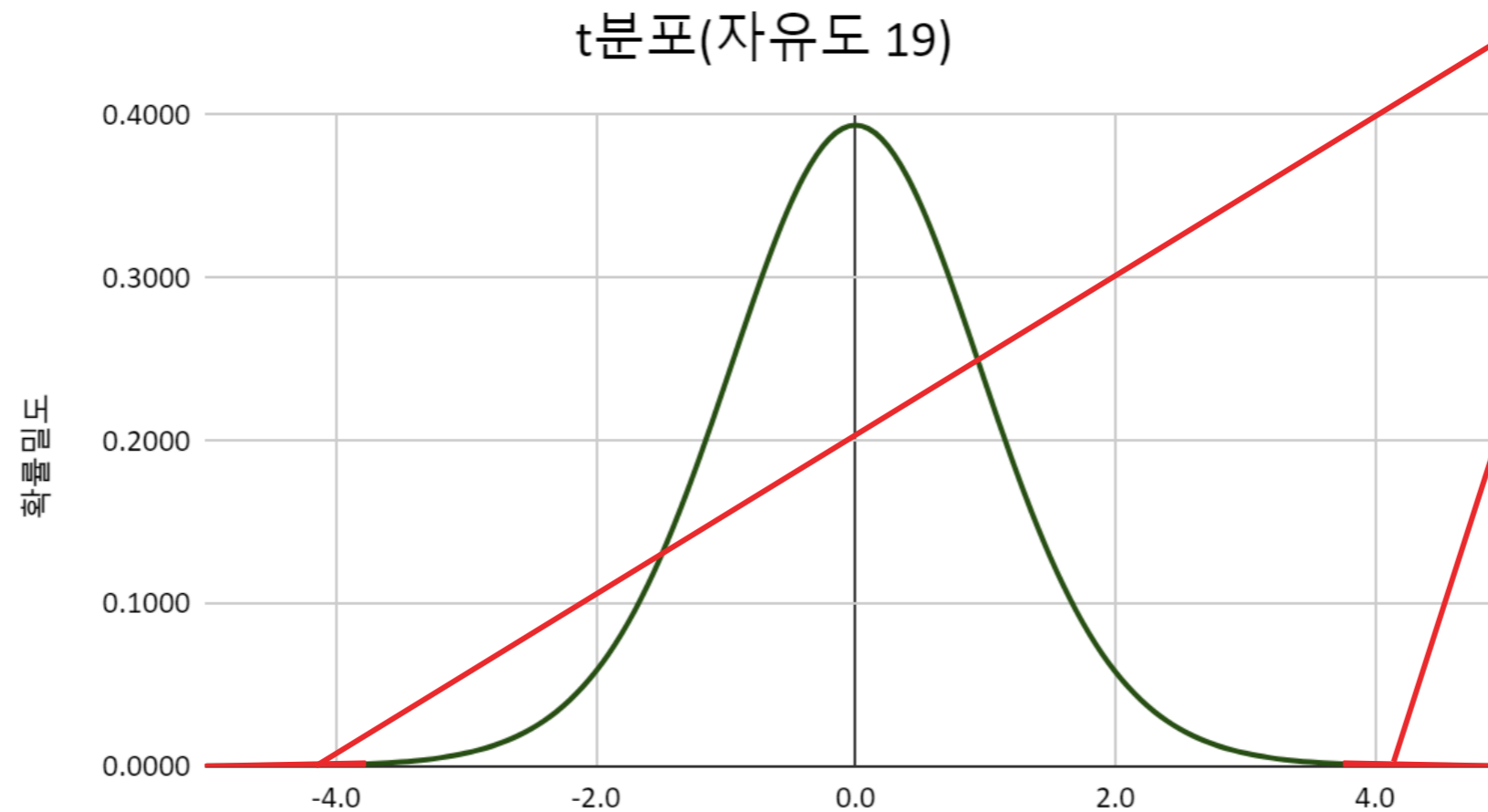
표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09



검정통계량이 나눈 세 구간 중 평균(0)이 속하지 않은 구간의 면적이 유의확률

유의확률

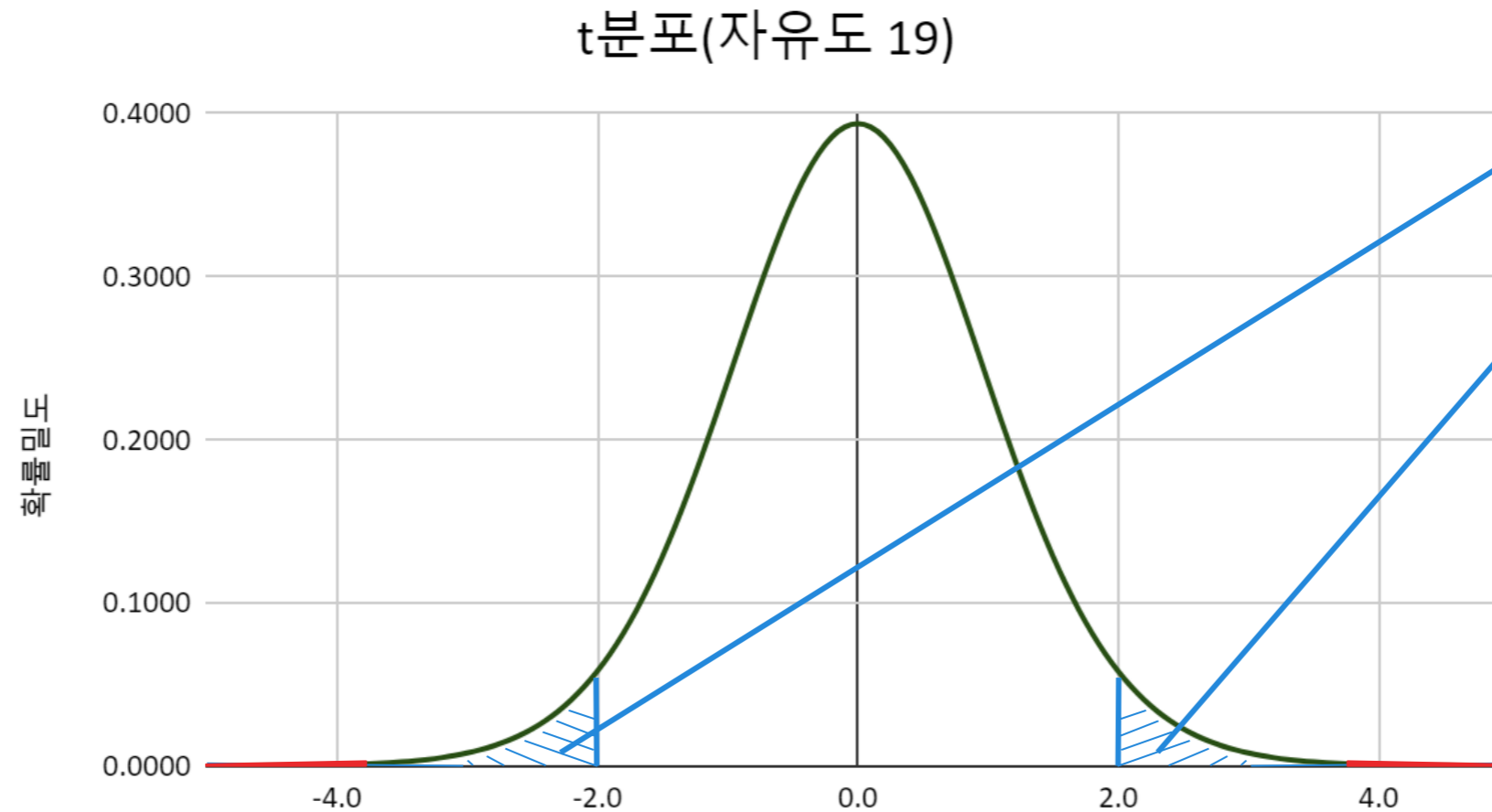
표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09



검정통계량이 나눈 세 구간 중 평균(0)이 속하지 않은 구간의 면적이 유의확률

유의수준

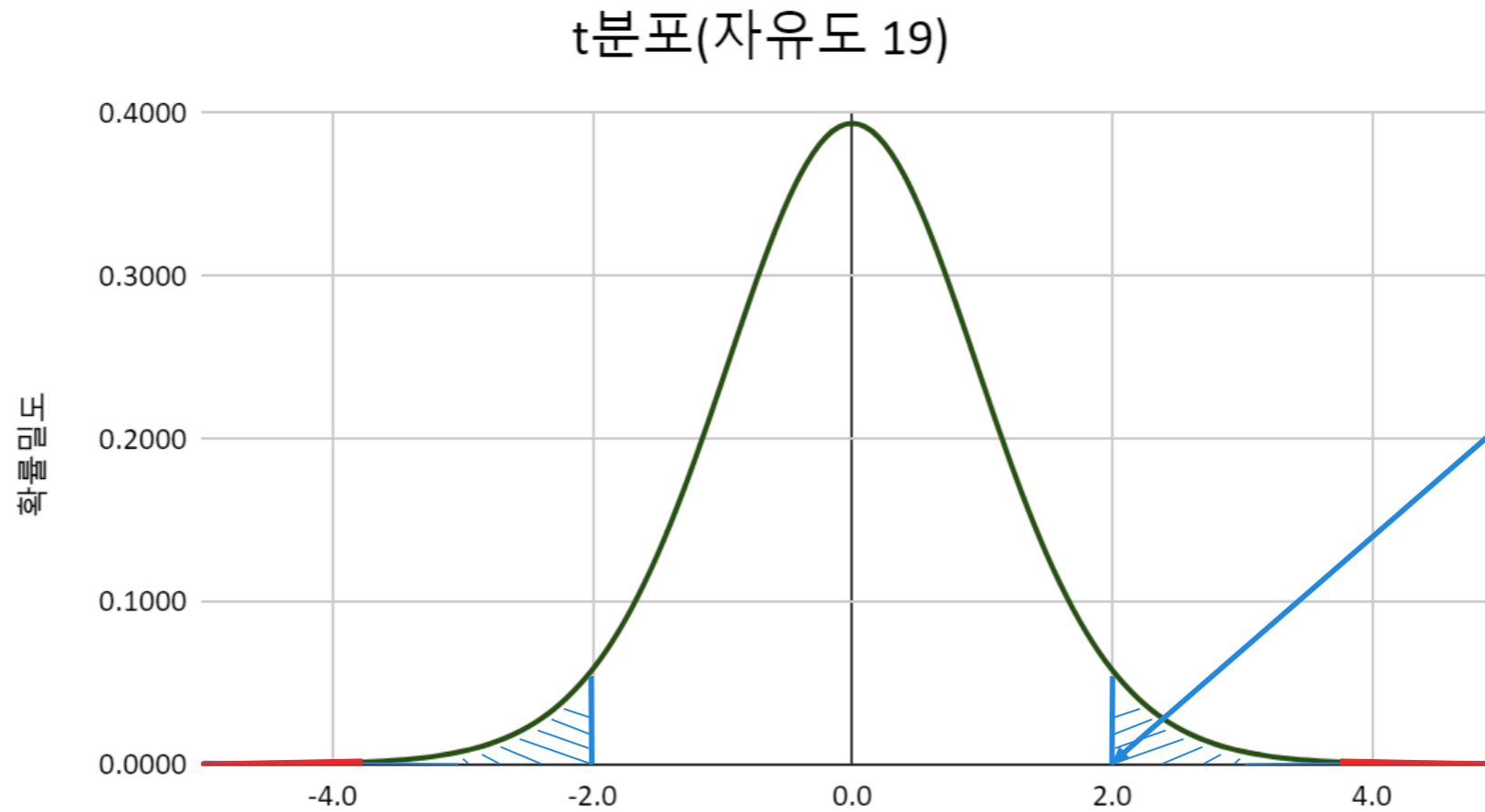
표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09



“저온숙성 전후 딸기의 당도 차이는 0이다”라는 가설을 95% 신뢰한다면 유의수준은 5%

유의수준

표본크기	사전 당도 평균	사후 당도 평균	차이 평균	자유도	차이 분산	차이 표준편차	표준오차	검정통계량	유의확률	유의수준	임계값
20	12.3	12.5	0.2	19	0.1	0.2	0.05	3.78	0.001	0.050	2.09



“저온숙성 전후 딸기의 당도 차이는 0이다”라는 가설을 95% 신뢰한다면 유의수준은 5%

가설검정

연구주제

저온숙성을 하고 나면 딸기 당도가 향상된다.

귀무가설

저온숙성 전후 딸기의 당도 차이 모평균은 0이다.

확률분포

귀무가설이 옳다면, 저온숙성 전후 딸기 당도의 차이를 t분포로 모델링

검정통계량

3.78

유의확률

0.001

유의수준

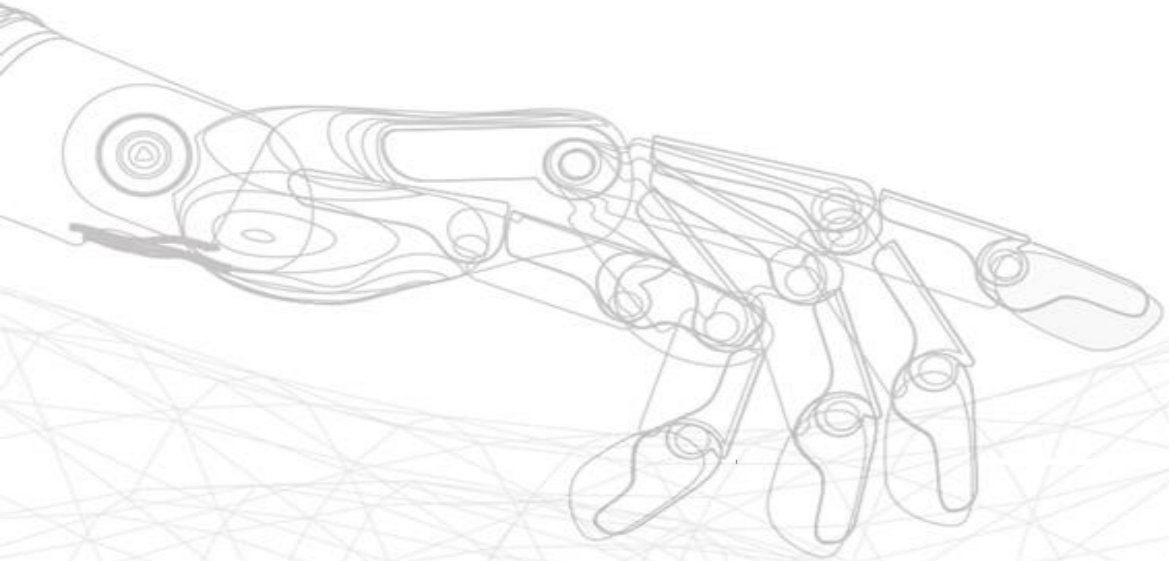
0.05(5%)

가설검정

“유의확률 < 유의수준” 이므로 귀무가설 기각

결론

귀무가설이 기각되었으므로 대립가설 채택



감사합니다

www.datadata.link

