

실험계획법 실습 - 4주차

## Chapter2 Basic Statistical Methods



황금비율의 소맥잔! 어떠신가요? 주목받고 있는 소맥잔



# 1. Models for the Data and T-Test

- Model(2.23) is called a means model.

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2 \\ j = 1, 2, \dots, n_i \end{cases}$$

$$y_{ij} = \mu_i + \varepsilon_{ij} \begin{cases} i = 1, 2 \\ j = 1, 2, \dots, n_i \end{cases} \quad \mu_i = \mu + \tau_i$$

- $y_{ij}$ : Response of j-th observation on the i-th treatment.
- $\mu$ : A common effect for the whole experiment.
- $\tau_i$ : The i-th treatment effect.
- $\varepsilon_{ij}$ : Random error in the j-th observation on the i-th treatment,

**following  $NID(o, \sigma^2)$**  with common variance  $\sigma^2$ .

**$NID$**  = Normally and Identically Distributed

# 1. Models for the Data and T-Test

## ● EX1) Dataset: Table2.1

■ TABLE 2.1

Tension Bond Strength Data for the Portland Cement Formulation Experiment

	Modified Mortar	Unmodified Mortar
$j$	$y_{1j}$	$y_{2j}$
1	16.85	16.62
2	16.40	16.75
3	17.21	17.37
4	16.35	17.12
5	16.52	16.98
6	17.04	16.87
7	16.96	17.34
8	17.15	17.02
9	16.59	17.08
10	16.57	17.27

Table 2.1  
© John Wiley & Sons, Inc. All rights reserved.

$y_{ij}$ : Tension의 Strength(강도)

Mortar: 시멘트와 모래를 물로 반죽  
이 실험에서 원하는 건?

여기서의 Model?

$$1) y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \varepsilon_{ij} \sim NID(o, \sigma^2)$$

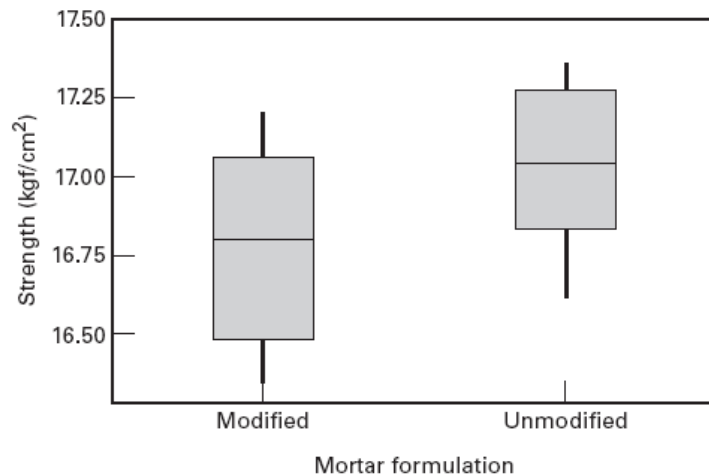
$$2) y_{ij} = \mu_i + \varepsilon_{ij}, \varepsilon_{ij} \sim NID(o, \sigma^2)$$

단,  $i = 1, 2$   $j = 1, 2, 3, 4, \dots, 10$

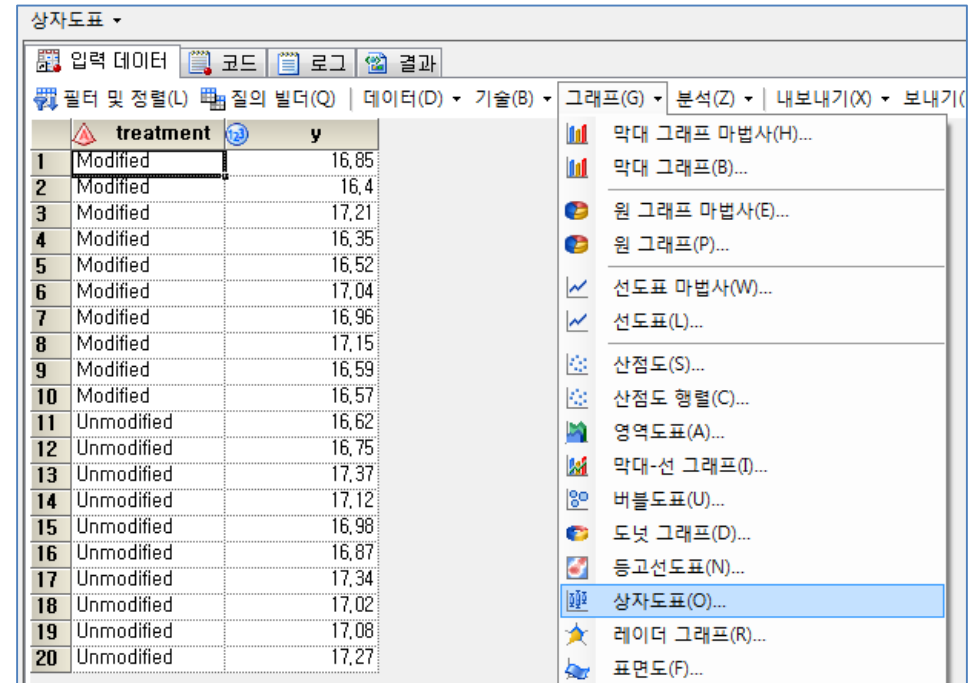
Model에 대한 식 및 가정이 없을 시 **감점**처리!

# 1. Models for the Data and T-Test

## ● Box Plots

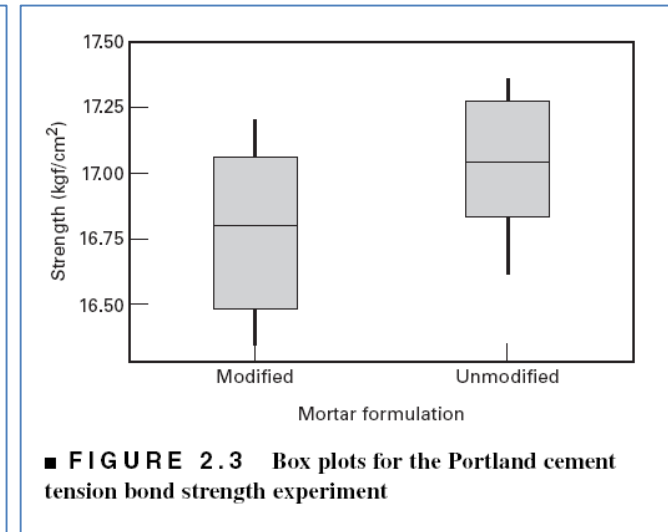
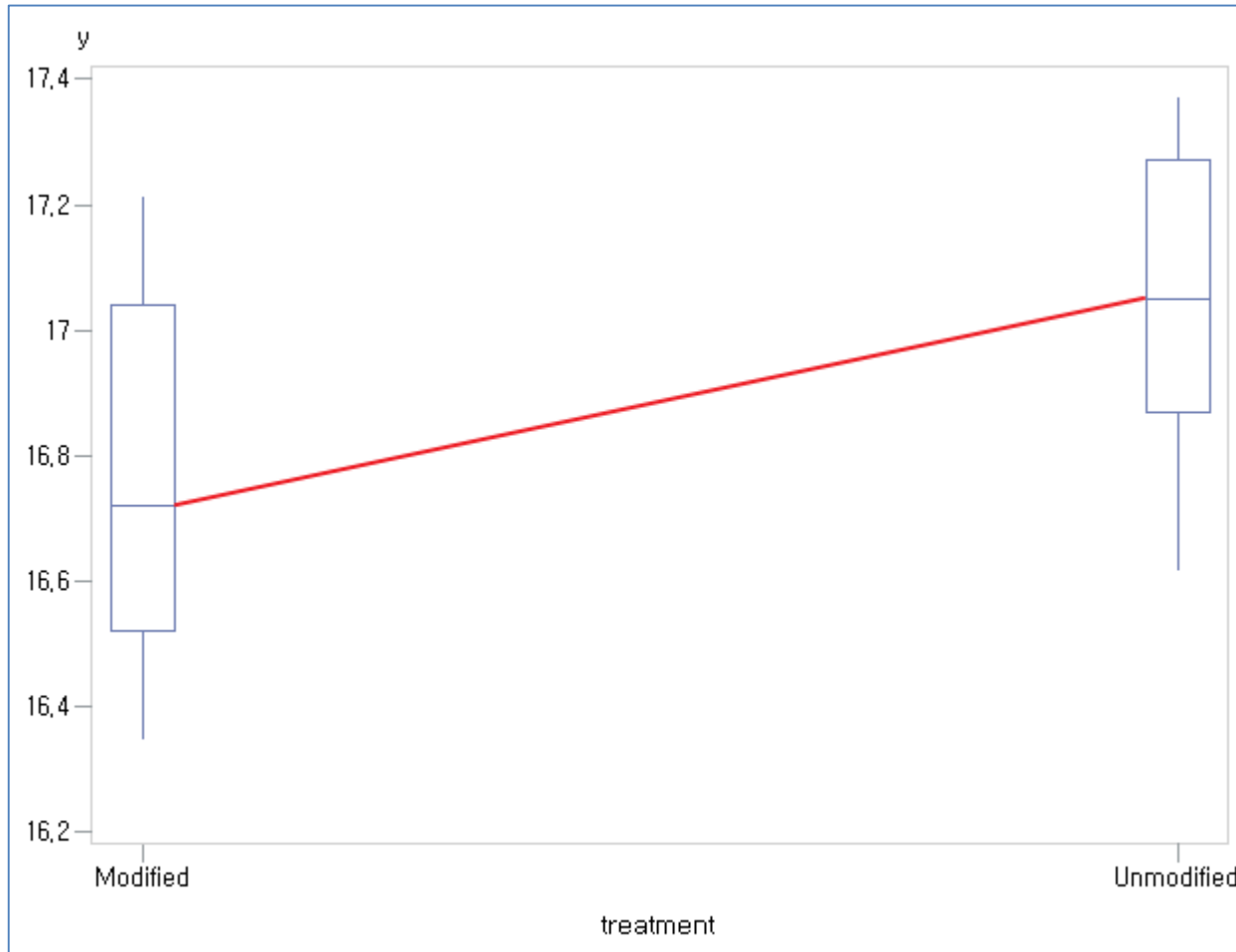


■ FIGURE 2.3 Box plots for the Portland cement tension bond strength experiment



# 1. Models for the Data and T-Test

## ● Result



# 1. Models for the Data and T-Test

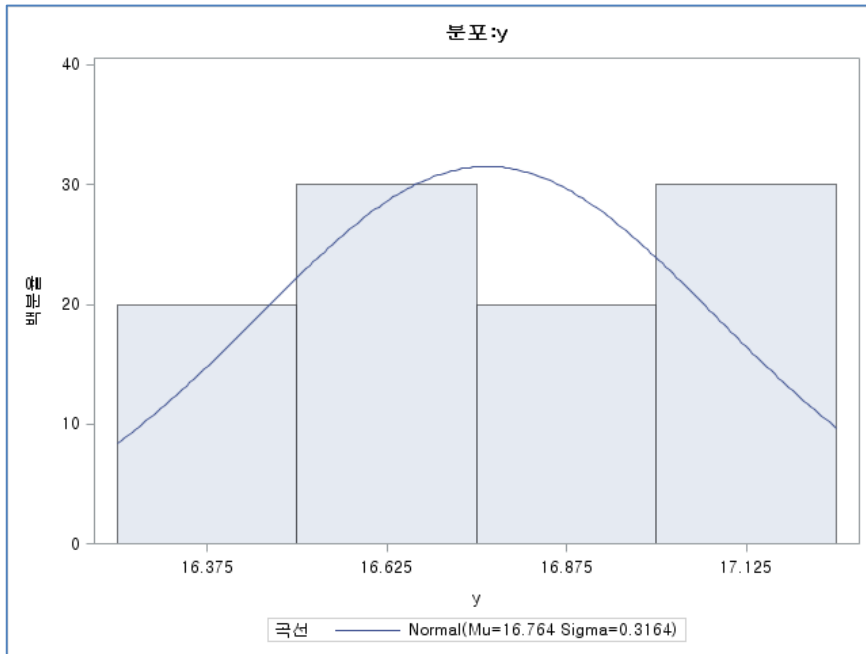
## ● Histogram(Treatment 효과 제외한 히스토그램 그려보기)

	treatment	y
1	Modified	16,85
2	Modified	16,4
3	Modified	17,21
4	Modified	16,35
5	Modified	16,52
6	Modified	17,04
7	Modified	16,96
8	Modified	17,15
9	Modified	16,59
10	Modified	16,57
11	Unmodified	16,62
12	Unmodified	16,75
13	Unmodified	17,37
14	Unmodified	17,12
15	Unmodified	16,98
16	Unmodified	16,87
17	Unmodified	17,34
18	Unmodified	17,02
19	Unmodified	17,08
20	Unmodified	17,27

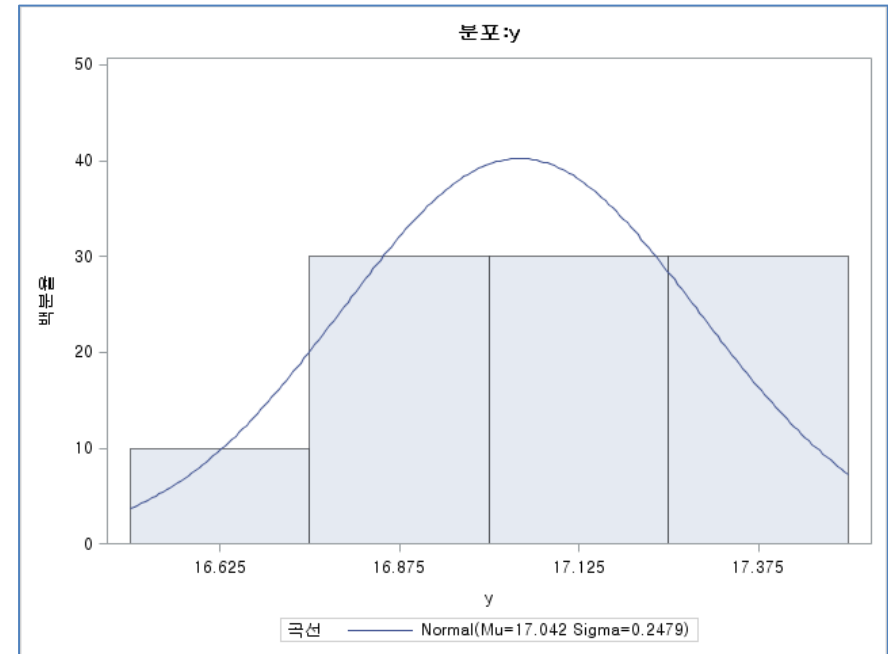
데이터	데이터	
분포	데이터 소스:	
요약	작업 필터:	
정규		
로그 정규		
지수	할당할 변수(A):	작업 역할(T):
와이블	이름	분석변수
베타	treatment	y
감마	y	그룹 분석변수
커널		빈도변수(제한: 1개)
도표		상대 가중값 변수(제한: 1개)
모양		분류변수(제한: 2개)
인셋		
테이블		
제목		
속성		

# 1. Models for the Data and T-Test

## ● Result



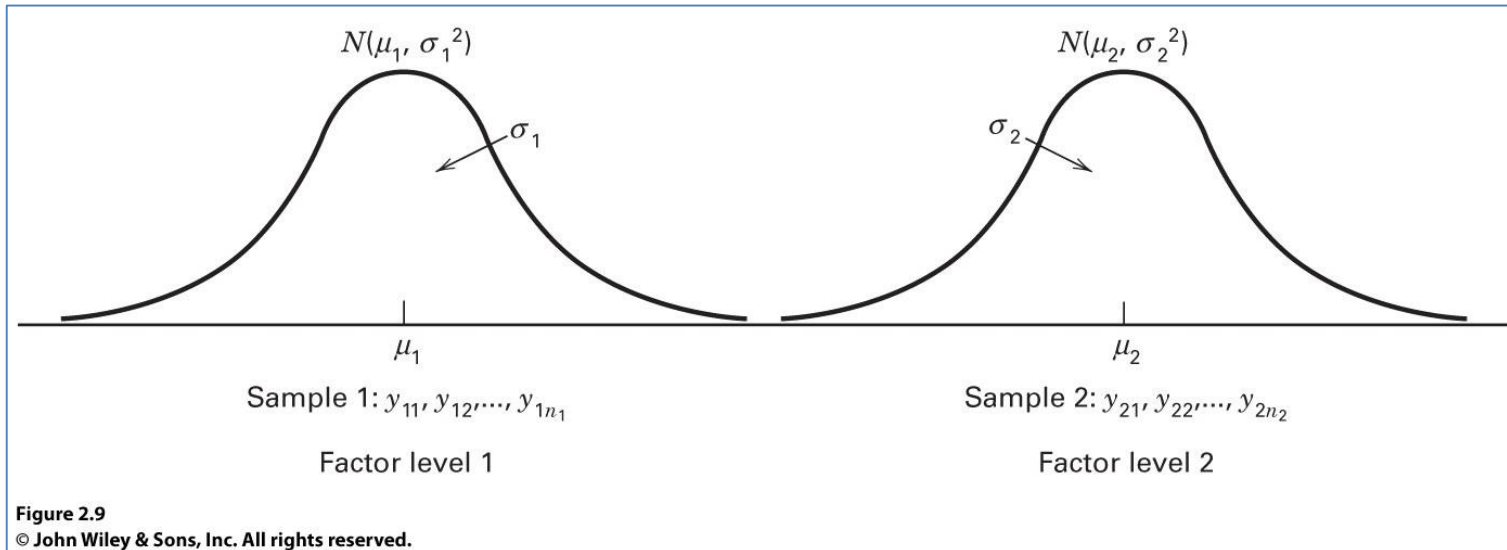
Modified



Unmodified

## 2. The Hypothesis Testing Framework

### ● T-Test

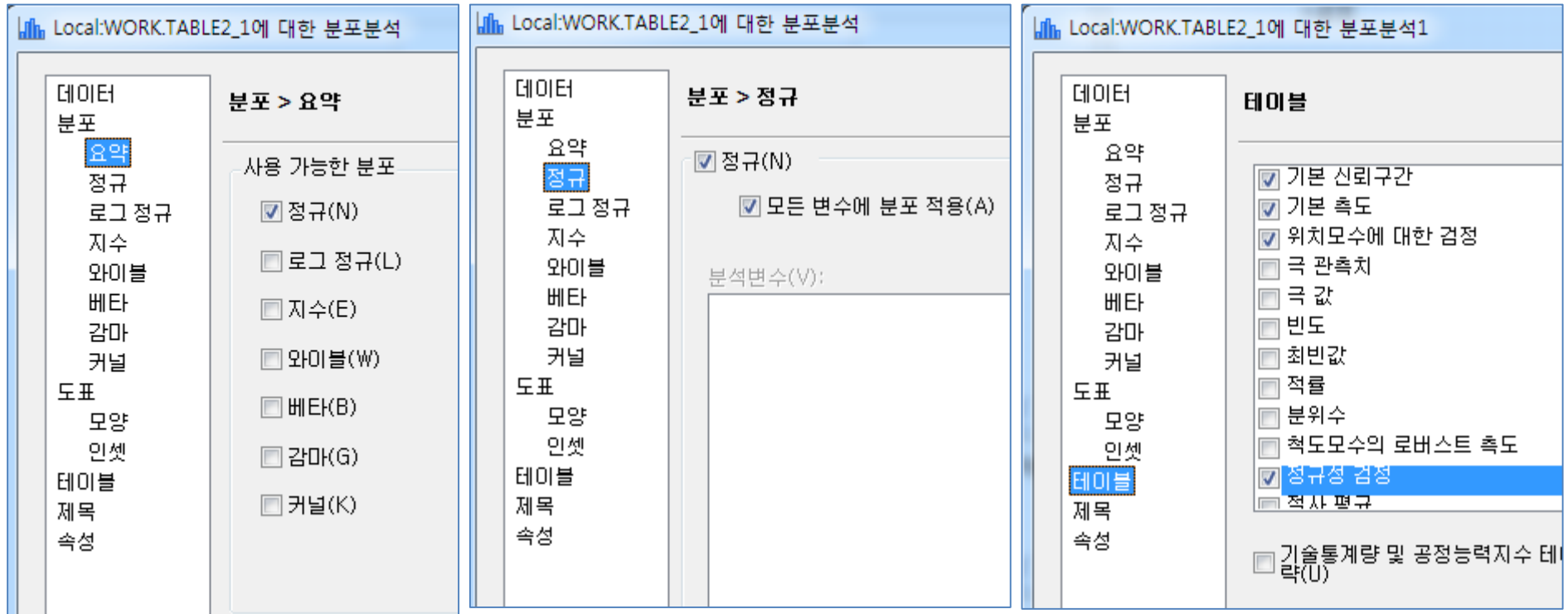


- Sampling form **a normal distribution**.
- How hypotheses ( **$\mu$** 에 대해서)  
 $H_0$ : (직접 적어보세요.)  
 $H_1$ : (직접 적어보세요.)



## 2. The Hypothesis Testing Framework

- EX2) Table2.1 treatment에 따라 Normal distribution을 만족?



- Shapiro-Wilk 통계량 확인!! ([http://en.wikipedia.org/wiki/Shapiro%E2%80%93Wilk\\_test](http://en.wikipedia.org/wiki/Shapiro%E2%80%93Wilk_test))
- 그 외 Kolmogorov-Smirnov, Cramer-von Mises, Anderson-Darling이 존재.

## 2. The Hypothesis Testing Framework

### ● Result

분포분석: y  
UNIVARIATE 프로시저  
변수: y  
treatment=Modified

기본 통계 측도			
위치 측도		변이 측도	
평균	16.76400	표준 편차	0.31645
중위수	16.72000	분산	0.10014
최빈값	.	범위	0.86000
		사분위 범위	0.52000

정규성 검정			
검정	통계량	p 값	
Shapiro-Wilk	W	0.918633	Pr < W 0.3457
Kolmogorov-Smirnov	D	0.208791	Pr > D >0.1500
Cramer-von Mises	W-Sq	0.059315	Pr > W-Sq >0.2500
Anderson-Darling	A-Sq	0.352804	Pr > A-Sq >0.2500

분포분석: y  
UNIVARIATE 프로시저  
변수: y  
treatment=Unmodified

기본 통계 측도			
위치 측도		변이 측도	
평균	17.04200	표준 편차	0.24792
중위수	17.05000	분산	0.06146
최빈값	.	범위	0.75000
		사분위 범위	0.40000

정규성 검정			
검정	통계량	p 값	
Shapiro-Wilk	W	0.962619	Pr < W 0.8153
Kolmogorov-Smirnov	D	0.121127	Pr > D >0.1500
Cramer-von Mises	W-Sq	0.021149	Pr > W-Sq >0.2500
Anderson-Darling	A-Sq	0.169975	Pr > A-Sq >0.2500

- 정규성을 만족?(유의수준 0.05 기준)
- 히스토그램 도표로 확인이 힘들 경우 Q-Q 도표로도 확인 가능!

## 2. The Hypothesis Testing Framework

- 그렇다면? Treatment에 대한 효과가 있을까?

The screenshot displays the Minitab interface. On the left, a data table with columns 'treatment' and 'y' is shown. The 'treatment' column contains 'Modified' for rows 1-10 and 'Unmodified' for rows 11-20. The 'y' column contains numerical values. On the right, the 't-검정 유형' (t-test type) dialog box is open, showing options for '이표본(T)' (Two-sample), '쌍체(P)' (Paired), and '일표본(O)' (One-sample). The '이표본(T)' option is selected, and a normal distribution curve is shown.

	treatment	y
1	Modified	16.85
2	Modified	16.4
3	Modified	17.21
4	Modified	16.35
5	Modified	16.52
6	Modified	17.04
7	Modified	16.96
8	Modified	17.15
9	Modified	16.59
10	Modified	16.57
11	Unmodified	16.62
12	Unmodified	16.75
13	Unmodified	17.37
14	Unmodified	17.12
15	Unmodified	16.98
16	Unmodified	16.87
17	Unmodified	17.34
18	Unmodified	17.02
19	Unmodified	17.08
20	Unmodified	17.27

- 여기서 어떤 검정 유형을 선택해야 하는지?(양측, 유의수준 0.05)

## 2. The Hypothesis Testing Framework

### ● Result

t-검정						
The TTEST Procedure						
Variable: y						
treatment	N	Mean	Std Dev	Std Err	Minimum	Maximum
Modified	10	16.7640	0.3164	0.1001	16.3500	17.2100
Unmodified	10	17.0420	0.2479	0.0784	16.6200	17.3700
Diff (1-2)		-0.2780	0.2843	0.1271		

treatment	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
Modified		16.7640	16.5376 16.9904	0.3164	0.2177 0.5777
Unmodified		17.0420	16.8647 17.2193	0.2479	0.1705 0.4526
Diff (1-2)	Pooled	-0.2780	-0.5451 -0.0109	0.2843	0.2148 0.4204
Diff (1-2)	Satterthwaite	-0.2780	-0.5462 -0.00983		

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	18	-2.19	0.0422
Satterthwaite	Unequal	17.025	-2.19	0.0430

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	9	9	1.63	0.4785

TABLE 2.2

Computer Output for the Two-Sample t-Test

Minitab

Two-sample T for Modified vs Unmodified

	N	Mean	Std. Dev.	SE Mean
Modified	10	16.764	0.316	0.10
Unmodified	10	17.042	0.248	0.078

Difference = mu (Modified) - mu (Unmodified)

Estimate for difference: -0.278000

95% CI for difference: (-0.545073, -0.010927)

T-Test of difference = 0 (vs not = ): T-Value = -2.19

P-Value = 0.042 DF = 18

Both use Pooled Std. Dev. = 0.2843

JMP t-test

Unmodified-Modified

Assuming equal variances

Difference	0.278000	t Ratio	2.186876
Std Err Dif	0.127122	DF	18
Upper CL Dif	0.545073	Prob>  t	0.0422
Lower CL Dif	0.010927	Prob> t	0.0211
Confidence	0.95	Prob< t	0.9789

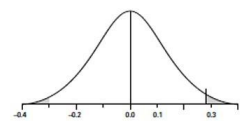


Table 2.2

© John Wiley & Sons, Inc. All rights reserved.

- 그렇다면 결론은?

# QUIZ

- 앞선 예제에 대해서 실험계획법을 다른 방법으로 풀 수 있지 않을까?

**이미 배웠습니다!!**



# REVIEW

- A와 B 다이어트 약의 효능을 비교하고자 한다.  
과 체중인 남성 10명을 Random하게 선택하고,  
2 집단에 5명씩 Random 하게 배정하였다.  
실험 전에 신체검사를 하여,  
집단 별로 과 체중의 평균을 비교한 결과  
유의한 차이는 없었다. 일정 기간 약을 복용한 후,  
체중 감소량(kg)을 측정하여 다음의 데이터를 얻었다.

다이어트 약		
OBS	A	B
1	4.3	6.4
2	5.8	6.6
3	5.1	5.9
4	5.5	6.0
5	4.5	6.1

Q. 어떤 다이어트 약이 체중감량에 더 효과가 있는가?

Model 및 가설의 형태를 세우고, 유의수준(0.05) 기준