

Monthly User Guide from JMP Korea

제 12호 (2018년 7월)

통계적 가설 검정(1)

- * 본 Guide 는 매월 세 번째 수요일에 발행됩니다
(2018년 7월호부터는 **JMP 14 Version** 기준입니다)
- ** Monthly User Guide 지난 호는 다음 Site 를 참조하세요(https://www.jmp.com/ko_kr/newsletters.html)
- *** 본 Guide 의 내용과 관련한 문의는 lkju.Shin@jmp.com 으로 연락 바랍니다

JMP 에서 가설 검정 Menu

통계적 가설 검정(또는 포괄적으로 '통계적 분석' 이라고 부르기도 합니다) 에 대한 JMP 기능[Menu] 는 아래와 같습니다.

본 Guide 에서는 JMP 를 활용한 통계적 가설 검정에 대한 개괄적으로 살펴보도록 하겠습니다.
(이번 호에 포함하지 못한 내용은 다음 기회에 포함하도록 하겠습니다)

가설 검정	JMP Menu	비고
Proportion Test	Analyze/Distribution ▼ Test Probabilities	Categorical Variable
1 Sample T Test	Analyze/Distribution ▼ Test Mean	
2 Sample T Test	Analyze / Fit Y by X ▼ Test Mean 또는 Analyze / Fit Y by X ▼ Means/Anova/Pooled T	Test Mean : Equal Variance 가정 안함 Means/Anova/Pooled T : Equal Variance 가정
One Way ANOVA	Analyze / Fit Y by X ▼ Means ANOVA	
회귀 분석(Regression)	Analyze / Fit Y by X	
Chi-Square Test	Analyze / Fit Y by X	Categorical Variable
Paired T Test	Analyze/Specialized Modeling / Matched Pairs	
Nonparametric Test (1 Sample)	Analyze/Distribution ▼ Test Mean 에서 Wilcoxon Signed Rank 선택	
Nonparametric Test (2 Sample & One way)	Analyze / Fit Y by X ▼ Nonparametric 에서 Wilcoxon Test 선택	
Two Way ANOVA	Analyze / Fit Model	Macros 에서 Full Factorial 선택

본 Report 의
설명 범위

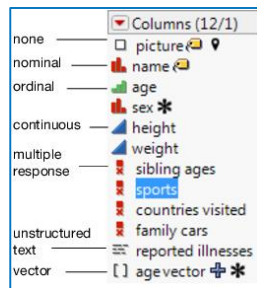
가설 검정을 하기 전에 확인해야 할 2가지

가설 검정을 하기 전에는 검정 대상이 되는 변수의 Modeling Type 과 분포(Distribution)를 반드시 확인해야 하는 데, 방법은 아래와 같습니다

Modeling Type 확인

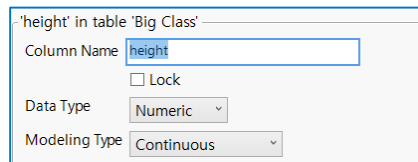
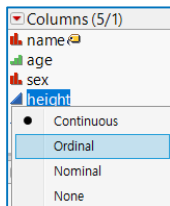
1. Modeling Type 확인 방법

- 1) Column Table 의 Column Head 명칭 왼쪽에 해당 변수의 Modeling Type 을 나타내는 Icon 이 구분 표시되어 있음
- 2) 이 중 Nominal(명목형), Ordinal(서열형, 순서형) 및 Continuous(연속형)는 반드시 사전 확인 필요



2. Modeling Type 의 변경

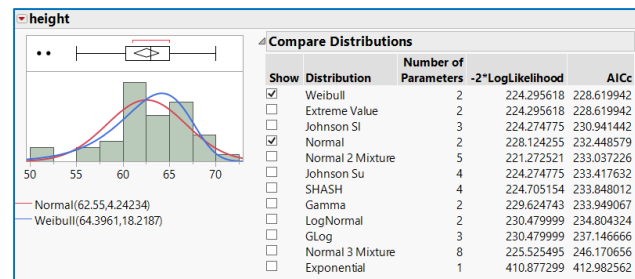
- 1) Modeling Type 을 표시하는 Icon 에서 오른쪽 마우스 Click 하여 변경 또는
- 2) Column Head 에서 오른쪽 마우스 Click → Column Info 에서 변경



분포(Distribution) 확인

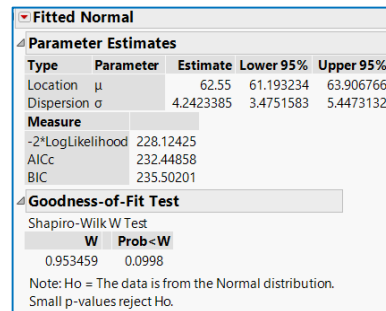
1. Analyze / Distribution 후에

- ▼ Continuous fit 또는 Discrete fit 하여 적합한 분포 확인



2. 정규성 검정(Normality Test) 하는 방법

- Analyze / Distribution 후에 ▼ Continuous fit ▼ Normal 클릭 후 Fitted Normal 에서 ▼ Goodness of Fit 선택



Proportion Test

1. Sample file : big class.jmp

2. Proportion Test(비율 분석)

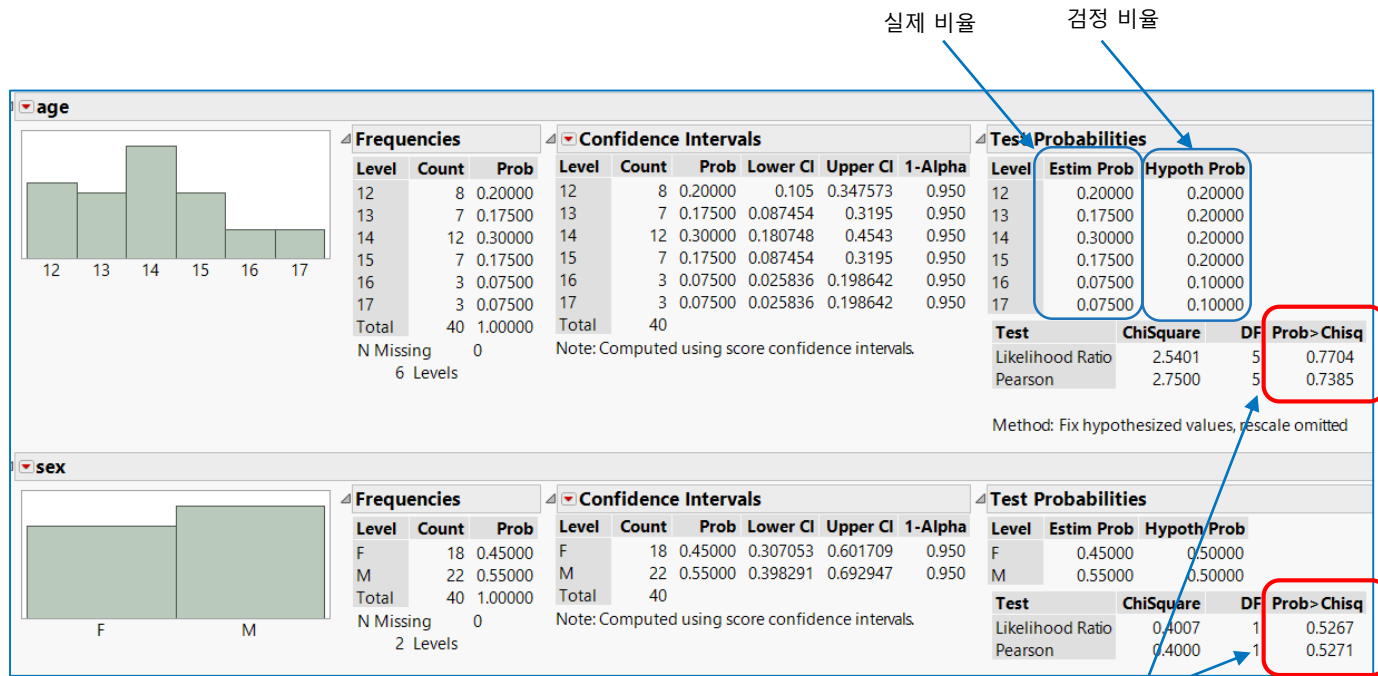
- 1) 비율 분석이므로 Categorical Test 만 해당
- 2) 나이 및 성별 Data 에 대해 Proportion Test 한다고 가정

3. Analyze / Distribution 에서 Age, Sex 선택 후 OK 클릭

4. ▼ Confidence Interval / 0.95 클릭
→ 신뢰 구간 확인

5. ▼ Test Probabilities 에서 비율 분석(검정)

- 1) Age
12세~15세까지는 각 20%(0.2) 및
16세~17세까지는 각 10%(0.1) 되는 지 검정
- 2) Sex(성별)
각 성별이 50%(0.5) 가 되는 지 검정



적어도 하나 이상의 샘플에 대한 비율이
검정 비율과 통계적으로 유의한 차이가 날 경우
이 값(P-Value)은 0.05 보다 작음
(유의 수준이 0.05 일 때)

1 Sample T Test

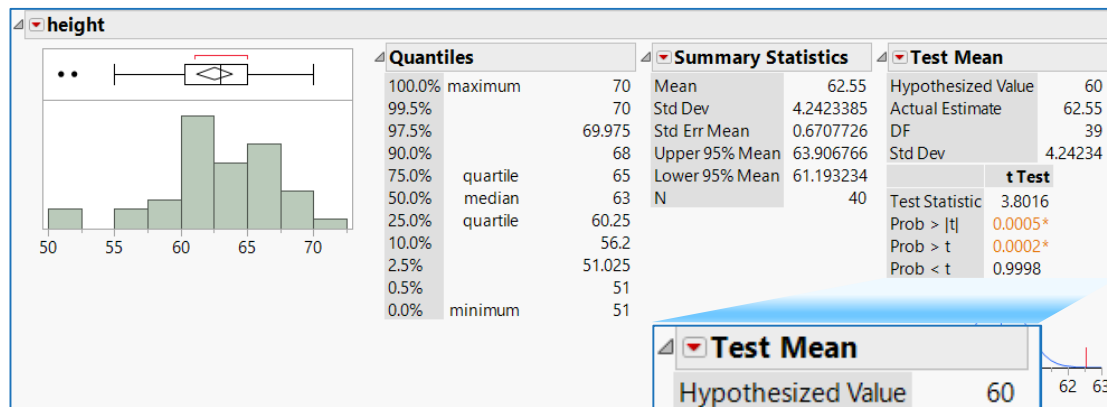
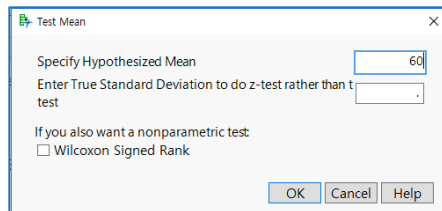
1. Sample file : big class.jmp

2. 1 Sample T Test

- 1) 연속형 데이터에 대한 평균치 검정
- 2) Height 에 대해 평균이 '60' 이라고 할 수 있는 지 검정한다고 가정

3. Analyze / Distribution 에서 Height 선택 후 OK 클릭

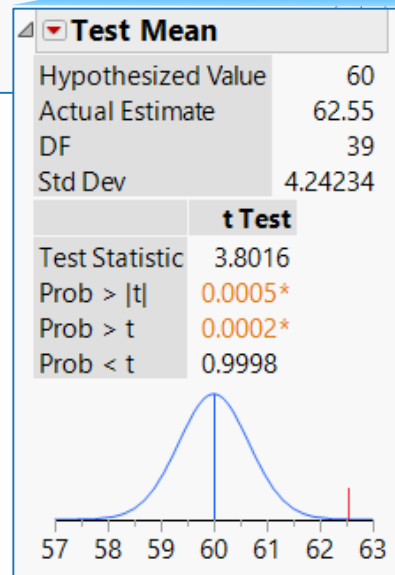
4. ▼ Test Mean 에서 'Specify Hypothesized Mean' 에 60 입력



평균이 60 이라는 귀무 가설을 기각한다는 뜻
양측 검정(평균이 60 이라고 할 수 없다는 뜻)

평균이 60 보다 작다는 귀무 가설을 기각한다는 뜻
단측 검정

평균이 60 보다 크다는 귀무 가설을 기각할 수 없다는 뜻
단측 검정



2 Sample T Test, One Way ANOVA & 회귀 분석

1. Sample file : big class.jmp

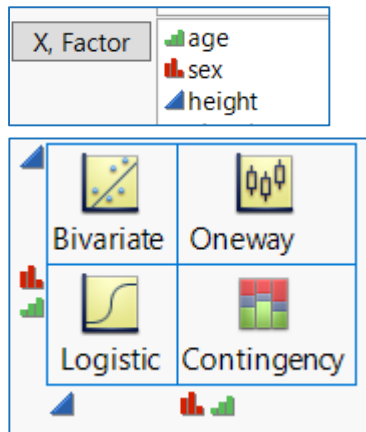
2. 가정

- 1) 반응치 : 몸무게(Weight)
- 2) 나이, 성별, 키 변수가 몸무게에 유의한 변수인 지를 검정하고자 함

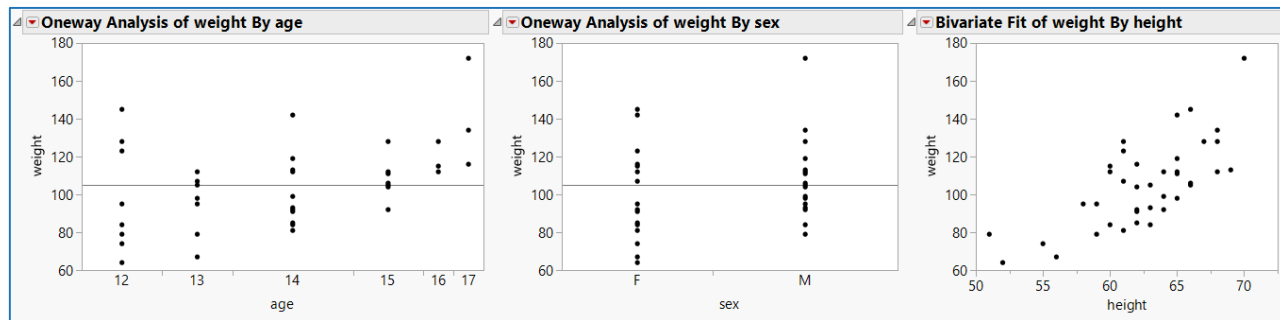
3. Analyze / fit Y by X 에 들어가서

- 1) Weight → Y
- 2) 나머지 3개 변수 → X
Height 선택 후 OK 클릭

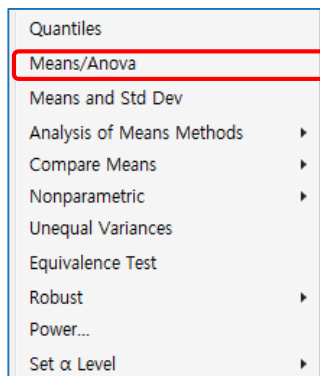
4. 3가지 X 변수의 Modeling Type 이 각각 다르므로, X 변수를 하나씩 선택해 보면 X 변수의 Modeling Type 에 따라 분석 방법이 다를 수 있음



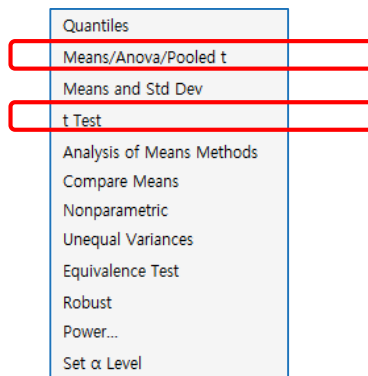
1. X 변수의 Modeling Type 에 따라 적절한 Graph 가 Display 됨



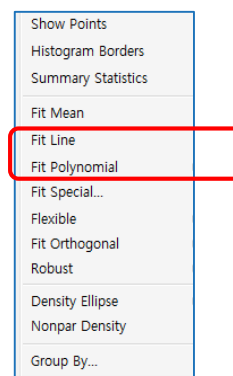
2. 각 변수명의 ▼ 를 Click 해 보면 하위 Menu 가 다를 수 있음(Modeling Type 에 따라 다른 분석을 하라는 뜻)



ANOVA 를 하라는 뜻



ANOVA 또는 T Test 를 하라는 뜻



회귀 분석(Regression) 을 하라는 뜻

2 Sample T Test, One Way ANOVA & 회귀 분석 : 2 Sample Test

1. Sample file : big class.jmp

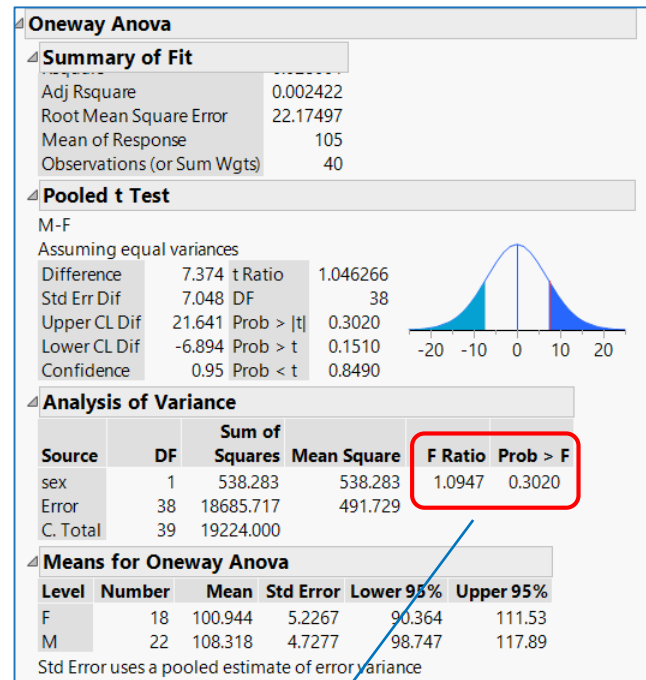
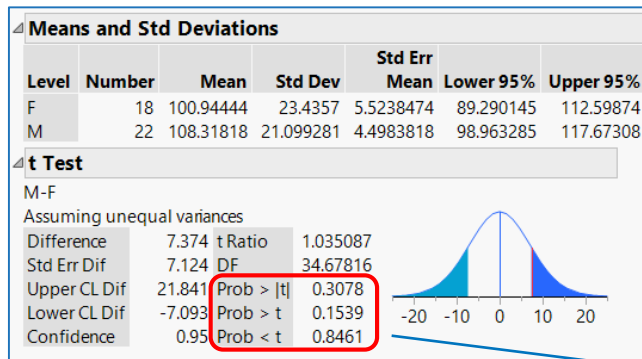
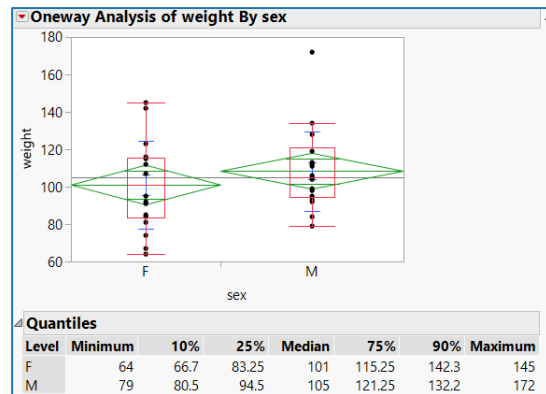
2. 가정

- 1) 반응치 : 몸무게(Weight)
- 2) 성별이 몸무게의 유의한 변수인지 검정

3. Alt key 를 누른 상태에서 ▼ 클릭
아래와 같이 네 가지 선택

Select Options and click OK

- ☒ Quantiles
- ☒ Means/Anova/Pooled t
- ☒ Means and Std Dev
- ☒ t Test



성별에 따라 몸무게의 유의한 차이가 있다고 말하기 어려움

2 Sample T Test, One Way ANOVA & 회귀 분석 : One Way ANOVA

1. Sample file : big class.jmp

2. 가정

- 1) 반응치 : 몸무게(Weight)
- 2) 나이에 따라 몸무게의 유의한 차이가 있는 지 검정

3. JMP 를 이용한 ANOVA 순서

1) ▼ Quantiles 및 ▼ Means and Std Deviations 를 Click

→ 수준(Level) 별 통계량 등 확인

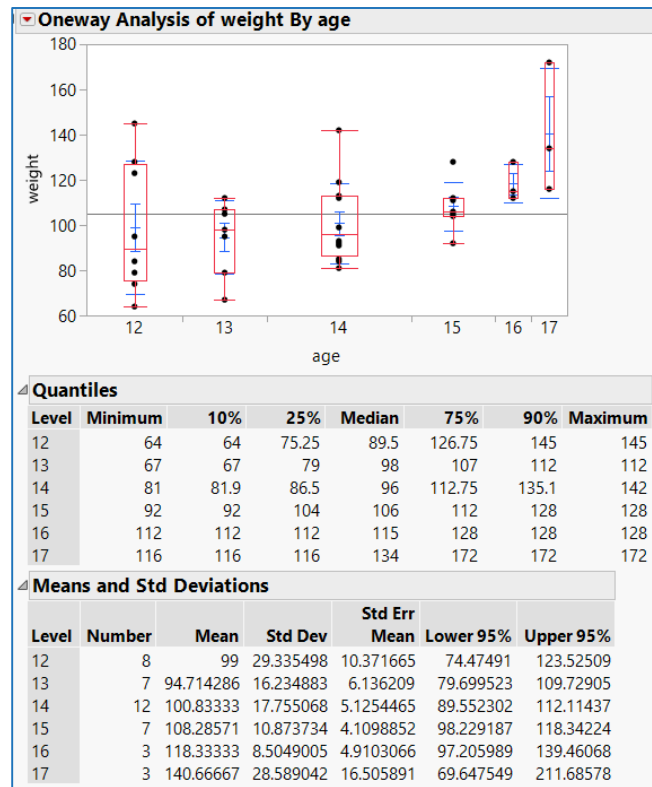
2) ▼ Means / ANOVA

→ ANOVA 실시

* ANOVA 에서의 유의 사항

수준(Level, 여기서는 나이) 별 비교 시 한 가지의 경우라도 통계적으로 유의하면 전체 ANOVA 결과는 유의한 것으로 결론 (여기서는 나이가 12~17세까지 6가지 Level 이므로 총 15가지의 Level 간 비교가 가능, 이 중 단 한 가지의 경우에서만 통계적으로 유의해도 전체적으로 유의하다고 결론)

→ 이에 대한 극복 방안은 다음 페이지



Oneway Anova

Summary of Fit

Rsquare	0.294535
Adj Rsquare	0.19079
Root Mean Square Error	19.97193
Mean of Response	105
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
age	5	5662.143	1132.43	2.8390	0.0301*
Error	34	13561.857	398.88		
C. Total	39	19224.000			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
12	8	99.000	7.061	84.65	113.35
13	7	94.714	7.549	79.37	110.06
14	12	100.833	5.765	89.12	112.55
15	7	108.286	7.549	92.94	123.63
16	3	118.333	11.531	94.90	141.77
17	3	140.667	11.531	117.23	164.10

Std Error uses a pooled estimate of error variance

2 Sample T Test, One Way ANOVA & 회귀 분석 : **One Way ANOVA**

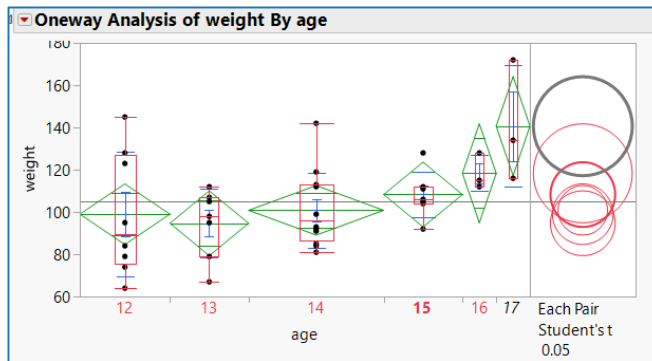
1. Sample file : big class.jmp

2. 가정

- 1) 반응치 : 몸무게(Weight)
- 2) 나이에 따라 몸무게의 유의한 차이가 있는 지 검정

3. JMP 를 이용한 ANOVA 순서

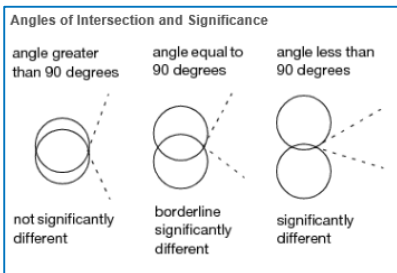
- 3) **▼ Compare Means**
/ Each Pair, Student T



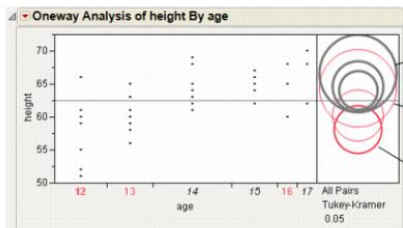
우측 동그라미 중 하나를 선택
→ 선택한 동그라미 선이 붉은 색으로 진하게 변함
(여기서는 15 살 선택)

15살과 유의한 차이가 있는 Level(여기서는 17세)은
검은색으로
유의한 차이가 없는 Level 은 선택한 동그라미
선과 같은 붉은 색으로 표시

<참고>



Highlighting Comparison Circles



Groups that are different from the selected group appear as thick gray circles.

Groups that are not different from the selected group appear as thin red circles.

The selected group appears as a thick red circle.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
17	13	45.95238	13.78194	17.9441	73.96066	0.0021*
17	12	41.66667	13.52106	14.1886	69.14477	0.0041*
17	14	39.83333	12.89183	13.6340	66.03268	0.0040*
17	15	32.38095	13.78194	4.3727	60.38923	0.0247*
16	13	23.61905	13.78194	-4.3892	51.62733	0.0957
17	16	22.33333	16.30702	-10.8065	55.47318	0.1798
16	12	19.33333	13.52106	-8.1448	46.81144	0.1619
16	14	17.50000	12.89183	-8.6993	43.69935	0.1836
15	13	13.57143	10.67545	-8.1237	35.26655	0.2123
16	15	10.04762	13.78194	-17.9607	38.05590	0.4710
15	12	9.28571	10.33646	-11.7205	30.29192	0.3753
15	14	7.45238	9.49855	-11.8510	26.75576	0.4381
14	13	6.11905	9.49855	-13.1843	25.42242	0.5238
12	13	4.28571	10.33646	-16.7205	25.29192	0.6810
14	12	1.83333	9.11590	-16.6924	20.35907	0.8418

17 세를 제외한
12~16 세 사이에는
유의한 차이가
거의 없음을 알 수 있음

2 Sample T Test, One Way ANOVA & 회귀 분석 : 회귀 분석

1. Sample file : big class.jmp

2. 가정

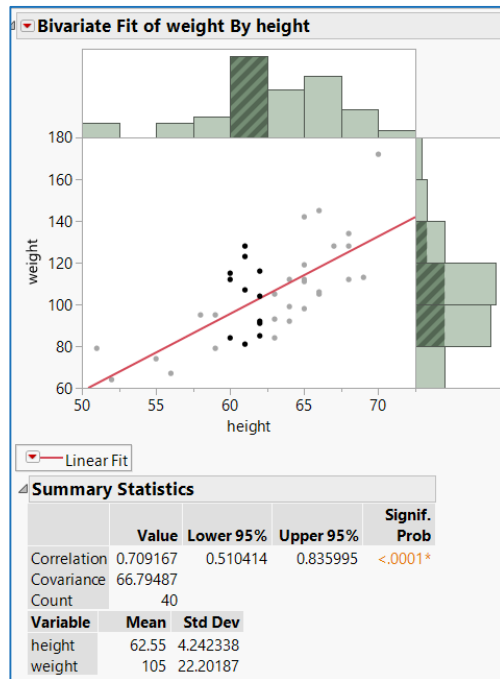
- 1) 반응치 : 몸무게(Weight)
- 2) 키(Height)에 따라 몸무게의 차이가 있는 지 검정
- 3) 산점도를 보면 키와 몸무게 간에는 선형 관계가 있는 것으로 추정됨.

3. ▼ Histogram Boarder 및

- ▼ Summary Statistics 클릭하여 그래프 및 통계량 확인

4. ▼ Fit Line에서 회귀 분석

- (만약, 산점도에서 2차 이상의 곡선 관계가 보일 경우에는
- ▼ Fit Polynomial 선택)



Linear Fit

weight = -127.1452 + 3.7113549*height

Summary of Fit

RSquare	0.502917
RSquare Adj	0.489836
Root Mean Square Error	15.85786
Mean of Response	105
Observations (or Sum Wgts)	40

Lack Of Fit

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	9668.079	9668.08	38.4460
Error	38	9555.921	251.47	Prob > F
C. Total	39	19224.000		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	-127.1452	37.52372	-3.39	0.0016*
height	3.7113549	0.598559	6.20	<.0001*

Chi-Square Test

1. Sample file : car poll.jmp

2. 분석 내용

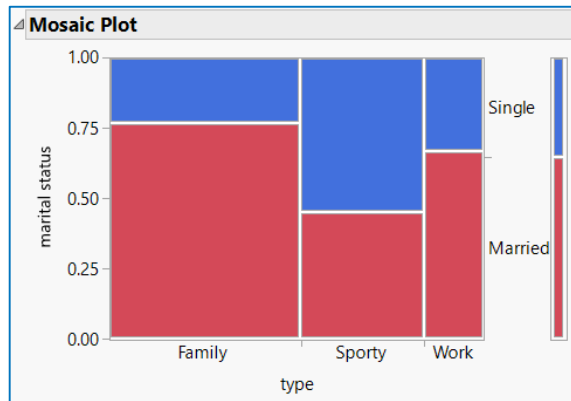
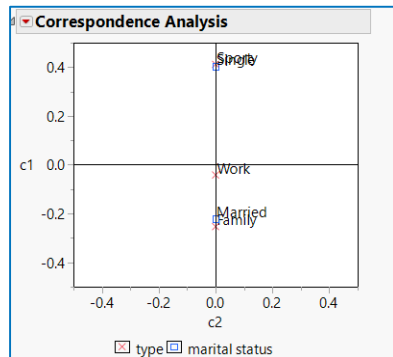
1) 결혼 여부(Marital Status)와
구매 목적별 차량 유형(type)과의
관련성 파악

2) Analyze / Fit Y by X 에서
Marital status → Y,
Type → X 입력

<참고>

대응 분석을 통해서도 관련성 확인 가능
Contingency Analysis 에서

▼ Correspondence analysis 클릭



Tests				
	N	DF	-LogLike	RSquare (U)
	303	2	13.382804	0.0680
Test	ChiSquare	Prob>ChiSq		
Likelihood Ratio	26.766	<.0001*		
Pearson	26.963	<.0001*		

Contingency Table	
<input checked="" type="checkbox"/>	Count
<input checked="" type="checkbox"/>	Total %
<input checked="" type="checkbox"/>	Col %
<input checked="" type="checkbox"/>	Row %
<input type="checkbox"/>	Expected
<input type="checkbox"/>	Deviation
<input type="checkbox"/>	Cell Chi Square
<input type="checkbox"/>	Col Cum
<input type="checkbox"/>	Col Cum %
<input type="checkbox"/>	Row Cum
<input type="checkbox"/>	Row Cum %

Contingency Table

▼를 Click 하면 Display 되는
통계량 추가 가능

Contingency Table				
		marital status		
Count		Married	Single	Total
Total %				
Col %				
Row %				
type	Family	119	36	155
		39.27	11.88	51.16
		60.71	33.64	
		76.77	23.23	
Sporty		45	55	100
		14.85	18.15	33.00
		22.96	51.40	
		45.00	55.00	
Work		32	16	48
		10.56	5.28	15.84
		16.33	14.95	
		66.67	33.33	
Total		196	107	303
		64.69	35.31	

Paired T Test

1. Sample file : Blood pressure by time.jmp

2. 가정

1) 짝을 이룬 Data(예 : 동일한 Sample 인데, 처리 전후의 data) 에 대한 유의차 분석

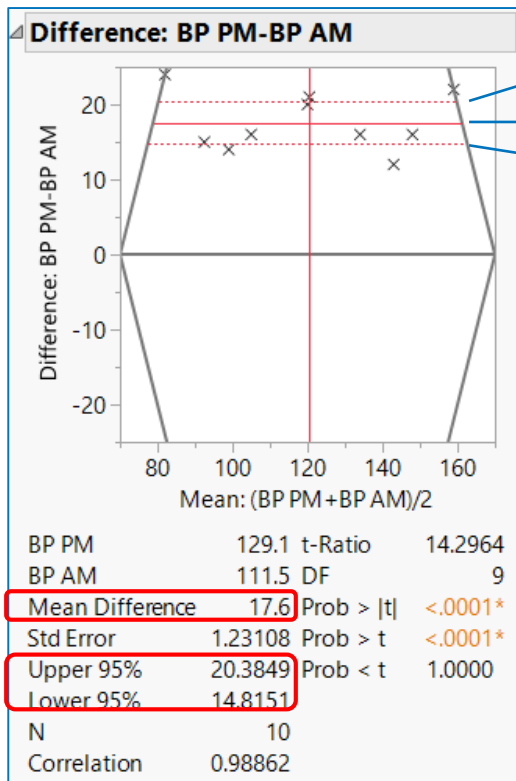
2) Sample Data

-동일인에 대한 혈압 측정 data
-오전/오후 혈압 간의 차이가 있는 지를 분석하고자 함

3. Analyze / Specialized Modeling
/ Matched Pairs

1) BP AM 과 BP PM 을 Y 값으로 선택

2) ▼ Reference Frame 선택



평균값

평균값의
95%
신뢰 구간

<참고>

차이 값에 대해 1 Sample T Test 실시한
결과와 동일함

