

전산통계 과제

4장 실습

컴퓨터소프트웨어 학부

2018008559

신상윤

추정

코드

```
data iris;
    infile '/home/u59322348/new_p/example-iris.txt' dlm = ',' firstobs=2;
    input Sepal_Length Sepal_Width Petal_Length Petal_Width Species
$;
run;
proc univariate data = iris cibasic alpha = 0.05;run;
proc univariate data = iris cibasic alpha = 0.01;run;
proc univariate data = iris cibasic alpha = 0.05;
    class Species;
run;
```

결과

Sepal_Length 전체데이터 신뢰수준 95%, 99%의 모평균, 모분산 구간추정

모수	추정값	95% 신뢰한계	
평균	5.84333	5.70973	5.97693
표준 편차	0.82807	0.74377	0.93408
분산	0.68569	0.55320	0.87250

모수	추정값	99% 신뢰한계	
평균	5.84333	5.66692	6.01975
표준 편차	0.82807	0.71977	0.97132
분산	0.68569	0.51807	0.94346

Sepal_Width

모수	추정값	95% 신뢰한계	
평균	3.05733	2.98701	3.12766
표준 편차	0.43587	0.39150	0.49167
분산	0.18998	0.15327	0.24174

모수	추정값	99% 신뢰한계	
평균	3.05733	2.96448	3.15019
표준 편차	0.43587	0.37886	0.51127
분산	0.18998	0.14354	0.26140

Petal_Length

모수	추정값	95% 신뢰한계	
평균	3.75800	3.47319	4.04281
표준 편차	1.76530	1.58560	1.99130
분산	3.11628	2.51412	3.96527

모수	추정값	99% 신뢰한계	
평균	3.75800	3.38192	4.13408
표준 편차	1.76530	1.53442	2.07069
분산	3.11628	2.35446	4.28775

Petal_Width

모수	추정값	95% 신뢰한계		모수	추정값	99% 신뢰한계	
평균	1,19933	1,07635	1,32231	평균	1,19933	1,03694	1,36172
표준 편차	0,76224	0,68464	0,85982	표준 편차	0,76224	0,66255	0,89410
분산	0,58101	0,46874	0,73929	분산	0,58101	0,43897	0,79942

Class(Species) 별로 각 변수의 95% 모평균, 모분산 구간추정

변수: Sepal_Length
Species = setosa

모수	추정값	95% 신뢰한계	
평균	5,93600	5,78931	6,08269
표준 편차	0,51617	0,43118	0,64322
분산	0,26643	0,18591	0,41373

변수: Sepal_Length
Species = versicol

모수	추정값	95% 신뢰한계	
평균	5,00600	4,90582	5,10618
표준 편차	0,35249	0,29445	0,43925
분산	0,12425	0,08670	0,19294

변수: Sepal_Length
Species = virginic

모수	추정값	95% 신뢰한계	
평균	6,58800	6,40729	6,76871
표준 편차	0,63588	0,53117	0,79239
분산	0,40434	0,28214	0,62788

변수: Sepal_Width
Species = setosa

모수	추정값	95% 신뢰한계	
평균	3,42800	3,32027	3,53573
표준 편차	0,37906	0,31665	0,47236
분산	0,14369	0,10026	0,22313

변수: Sepal_Width
Species = versicol

모수	추정값	95% 신뢰한계	
평균	2,77000	2,68082	2,85918
표준 편차	0,31380	0,26213	0,39103
분산	0,09847	0,06871	0,15291

변수: Sepal_Width
Species = virginic

모수	추정값	95% 신뢰한계	
평균	2,97400	2,88235	3,06565
표준 편차	0,32250	0,26939	0,40187
분산	0,10400	0,07257	0,16150

변수: Petal_Length
Species = setosa

모수	추정값	95% 신뢰한계	
평균	1,46200	1,41265	1,51135
표준 편차	0,17366	0,14507	0,21641
분산	0,03016	0,02104	0,04683

변수: Petal_Length
Species = versicol

모수	추정값	95% 신뢰한계	
평균	4,26000	4,12645	4,39355
표준 편차	0,46991	0,39253	0,58557
분산	0,22082	0,15408	0,34289

변수: Petal_Length
Species = virginic

모수	추정값	95% 신뢰한계	
평균	5,55200	5,39515	5,70885
표준 편차	0,55189	0,46102	0,68773
분산	0,30459	0,21254	0,47298

변수: Petal_Width
Species = setosa

모수	추정값	95% 신뢰한계	
평균	0,24600	0,21605	0,27595
표준 편차	0,10539	0,08803	0,13132
분산	0,01111	0,00775	0,01725

변수: Petal_Width
Species = versicol

모수	추정값	95% 신뢰한계	
평균	1.32600	1.26980	1.38220
표준 편차	0.19775	0.16519	0.24643
분산	0.03911	0.02729	0.06073

변수: Petal_Width
Species = virginic

모수	추정값	95% 신뢰한계	
평균	2.02600	1.94795	2.10405
표준 편차	0.27465	0.22942	0.34225
분산	0.07543	0.05264	0.11714

검정

코드

```
data iris;
    infile '/home/u59322348/new_p/example-iris.txt' dlm = ',' firstobs=2;
    input Sepal_Length Sepal_Width Petal_Length Petal_Width Species
$;
run;
proc univariate data = iris mu0=(6 3 4 1) alpha=0.05;
    var Sepal_Length Sepal_Width Petal_Length Petal_Width;
run;
proc means data = iris n mean var;
    var Sepal_Length Sepal_Width Petal_Length Petal_Width;
    output out = out_iris var(Sepal_Length Sepal_Width Petal_Length
Petal_Width) =
    v_Sepal_Length v_Sepal_Width v_Petal_Length v_Petal_Width;
run;
data out_iris;
    set out_iris;
    test1 = 149 * v_Sepal_Length / 0.2;
    test2 = 149 * v_Sepal_Width / 0.2;
    test3 = 149 * v_Petal_Length / 0.3;
```

```

test4 = 149 * v_Petal_Width / 0.06;
rejectL = quantile('chisquared',0.025,149);
rejectR = quantile('chisquared',0.975,149);

run;
proc print;run;

```

결과

각 평균 μ 에 대하여

$H_0 : \mu = x$, $H_1 : \mu \neq x$ 라 할 때

신뢰수준 95%에서 기각역은 $|T| \geq t_{\alpha/2}(149) = 1.97601$

각 검정통계량은

위치모수 검정: Mu0=6				
검정		통계량	p 값	
스튜던트의 t	t	-2,31717	Pr > t	0,0219
부호	M	-11	Pr >= M	0,0798
부호 순위	S	-1238,5	Pr >= S	0,0129

위치모수 검정: Mu0=3				
검정		통계량	p 값	
스튜던트의 t	t	1,611015	Pr > t	0,1093
부호	M	5	Pr >= M	0,4191
부호 순위	S	537,5	Pr >= S	0,1800

위치모수 검정: Mu0=4				
검정		통계량	p 값	
스튜던트의 t	t	-1,67897	Pr > t	0,0953
부호	M	11,5	Pr >= M	0,0673
부호 순위	S	-771,5	Pr >= S	0,1281

위치모수 검정: Mu0=1				
검정		통계량	p 값	
스튜던트의 t	t	3,202839	Pr > t	0,0017
부호	M	21,5	Pr >= M	0,0004
부호 순위	S	1345,5	Pr >= S	0,0061

이고, 각 경우에 검정 결과는

Sepal_Length : 기각역에 포함, $\alpha=0.05$ 하에서 H_0 를 기각한다. 따라서 Sepal_Length의 모평균을 6이라 할 수 없다.

Sepal_Width : 기각역에 포함되지 않음, $\alpha=0.05$ 하에서 H_0 를 채택한다. 따라서 Sepal_Width의 모평균을 3이라 할 수 있다.

Petal_Length : 기각역에 포함되지 않음, $\alpha=0.05$ 하에서 H_0 를 채택한다. 따라서 Petal_Length의 모평균을 4라 할 수 있다.

Petal_Width : 기각역에 포함, $\alpha=0.05$ 하에서 H_0 를 기각한다. 따라서 Petal_Width의 모평균을 1이라 할 수 없다.

각 분산 σ^2 에 대하여

$H_0 : \sigma^2 = x$, $H_1 : \sigma^2 \neq x$ 라 할 때

rejectL	rejectR
117,098	184,687

신뢰수준 95%에서 기각역은 $\chi^2 \geq 184.687$ or $\chi^2 \leq 117.098$

각 검정통계량은

test1	test2	test3	test4
510,842	141,535	1547,75	1442,83

이고, 각 경우에 검정 결과는

Sepal_Length : 기각역에 포함, $\alpha=0.05$ 하에서 H_0 를 기각한다. 따라서 Sepal_Length의 모분산을 0.2라 할 수 없다.

Sepal_Width : 기각역에 포함되지 않음, $\alpha=0.05$ 하에서 H_0 를 채택한다. 따라서 Sepal_Width의 모분산을 0.2라 할 수 있다.

Petal_Length : 기각역에 포함, $\alpha=0.05$ 하에서 H_0 를 기각한다. 따라서 Petal_Length의 모분산을 0.3이라 할 수 없다.

Petal_Width : 기각역에 포함, $\alpha=0.05$ 하에서 H_0 를 기각한다. 따라서 Petal_Width의 모분산을 0.06이라 할 수 없다.