

보건통계학특론

환경보건시스템학과 2학기 과목

담당: 이경무

튜터: 지석민

강의 개요

- ▶ 1. 무료 통계패키지인 R 프로그램을 이용하여 자료를 처리·분석하는 방법을 익힌다.
- ▶ 2. R을 이용한 자료의 처리·분석을 위해 기초적인 통계학적 지식을 익힌다.
- ▶ 3. 자신이 관심있는 자료를 선택한 후, 처리·분석하고, 결과를 정리·해석하는 과정을 수행함으로써 자료의 처리·분석에 자신감을 가진다.

강의 구성 (13개 강의)

- ▶ 1. R을 활용한 자료분석 입문
- ▶ 2. 양적자료의 정리
- ▶ 3. 이항분포와 정규분포
- ▶ 4. 표본분포와 추정
- ▶ 5. 가설검정의 이해
- ▶ 6. 평균 및 비율 차이에 대한 검정
- ▶ 7. 양적변수의 관련성에 대한 검정

강의 구성 (13개 강의)

- ▶ 8. 질적변수의 관련성에 대한 검정
 - ▶ 9. 비모수 통계
 - ▶ 10. 종적자료 분석
 - ▶ 11. 시계열자료 분석
 - ▶ 12. 공간자료 분석
 - ▶ 13. 메타분석
-
- ▶ 각 강의는 ppt 강의+R실습 강의
 - ▶ 학기 초(3/4 밤 8:00) 줌 실습: R설치 및 데이터 불러들이기

R, R studio: R실습

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

표본분포와 추정.R x 2강 실습_quantitative data.R x konehs3_mh_20191229.R* x 8강 실습_R x 1강 실습_qualitative data.R x 1강 참고 >>

```
1 #####
2 # 1강 실습: R을 이용한 자료분석 입문
3 #####
4
5 # 1) R 및 R studio 설치
6 # 2) 데이터 불러들이기
7 # 3) 데이터가 잘 불러들여 졌는지 확인하기
8 # 4) 질적 자료의 정리: 표, 그림
9
10 |
11
12 #데이터 불러들이기
13 ex=read.csv('d:/movie.csv')
14
15 #데이터 구조 확인
16 str(ex)
17
18 #변수의 형태 확인
19 typeof('grade')
20
21 #변수명 확인
22 names(ex)
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
```

Console

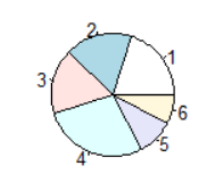
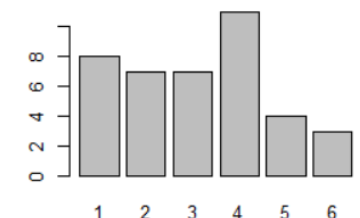
```
2 7 0.175 13 0.375
3 7 0.175 22 0.550
4 11 0.275 33 0.825
5 4 0.100 37 0.925
6 3 0.075 40 1.000
>
>
> #막대그림
> barplot(table(ex$choice_movie))
> barplot(table(ex$choice_movie), main='선호영화 막대그림', col=c('blue', 'skyblue', 'skyblue3', 'purple', 'red', 'yellow'))
>
> #원그림
> pie(table(ex$choice_movie))
> pie(table(ex$choice_movie), main='선호영화 Pie Chart', col=c('blue', 'skyblue', 'skyblue3', 'purple', 'red', 'yellow'))
>
```

Environment

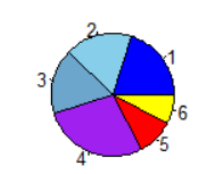
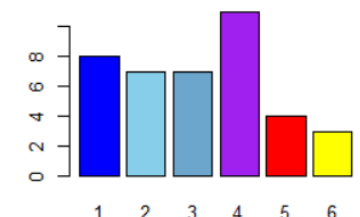
Object	Attributes
c	List of 6
data	20 obs. of 3 variables
data2	60 obs. of 2 variables
ex	40 obs. of 5 variables

Files Plots Packages Help Viewer

Zoom Export Publish



선호영화 막대그림



선호영화 Pie Chart

```
#####
# 2강 실습: 양적자료의 정리
#####

# 1) 도수분포표, 히스토그램
# 2) 줄기-잎 그림
# 3) 상자그림
# 4) 산포도

# 데이터 불러들이기
exam=read.csv('d:/score.csv')

# 변수명 확인
names(exam)

#갯수와 변수의 수 확인
dim(exam)

# 도수분포표의 계급 지정
a=cut(exam$final, breaks=c(21.5,33.5,45.5,57.5, 69.5, 81.5,
93.5))

# 도수분포표
table(a)

# 도수, %, 누적도수 누적 %
cbind(table(a), table(a)/length(a), cumsum(table(a)),
cumsum(table(a)/length(a)))
```

```
# 히스토그램
hist(exam$final)
hist(exam$final, breaks=8) # 구간의 개수 지정

# 계급 지정 시
hist(exam$final,breaks=c(21.5,33.5,45.5,57.5, 69.5, 81.5, 93.5))

# 상대도수를 Y 축에
hist(exam$final, probability=TRUE)

# 색깔 추가
hist(exam$final, probability=TRUE, col=c('blue', 'skyblue', 'skyblue3', 'purple',
'red', 'yellow'))

# 제목 달기
hist(exam$final, probability=TRUE, col=c('blue', 'skyblue', 'skyblue3', 'purple',
'red', 'yellow'), xlab="기말점수", ylab="상대도수", main="기말점수 히스토그램")

# 분포선 추가
lines(density(exam$final), col='black')

# 줄기-잎 그림
stem(exam$mid)

# 상자그림
boxplot(exam$mid, exam$final)

# 산점도
plot(exam$mid, exam$final)
```

평가방법

- ▶ 매 강의의 R 실습과제 13회 (20%)
- ▶ 중간시험(이론+R실습) (30%)
- ▶ 기말과제: 자료 선정, 처리·분석, 결과의 정리·해석 (R script 포함) (40%)
- ▶ 학습참여도 등 (10%)

참고자료

- ▶ 1. 정미숙, Excel과 R을 이용한 통계학. 자유아카데미. 2011.
- ▶ 2. 안재형. R을 이용한 누구나 하는 통계분석. 한나래. 2011.
- ▶ 3. 배정민. 닥터 배의 술술 보건의학통계. 한나래. 2012.
- ▶ 4. 서울대학교 자연과학대학 통계학과. 현대통계학. 2013.

- ▶ <참고사이트>
- ▶ 1. R 웹사이트: www.r-project.org
- ▶ 2. R을 이용한 기초 통계 처리:
<http://www.kocw.net/home/search/kemView.do?kemId=332938>
- ▶ 3. 자료처리 관련 자료: <http://plaza.snu.ac.kr/~hokim/> (-> 열린강의실)

수강할까 말까?

- ▶ 업무관련 자료 등, 직접 분석해 보고 싶다. -> 수강 추천
- ▶ 학위논문을 작성해 보고 싶다. -> 수강 추천
- ▶ 통계학 강의 수강한 적 없다. -> 수강 가능

감사합니다.