

## 3강. 통계계산(1)

◆ 담당교수 : 심송용 교수

### ■ 주요용어

용어	해설
모의실험	실제상황을 흉내(imitation)내는 가상 실험으로 주로 컴퓨터를 사용하는 실험이 많다.
분위수	q보다 같거나 작을 확률이 p일 때 q를 q번째 분위수 (q-th quantile)이라고 한다.
분포함수	x보다 같거나 작을 확률을 F(x)로 표현하고 이를 분포 함수라고 한다.
확률밀도함수	연속인 확률변수에서 임의의 구간 (a,b)에 속할 확률이 함수 f(x)를 피적분함수로 하는 (a,b) 사이의 적분일 때, 즉, $P[a < X < b] = \int_a^b f(x)dx$ 일 때 f(x)를 확률밀도함수라고 한다.

### ■ 연습문제

1. 5개의 원소를 가진 모집단에서 비복원으로 3개를 추출하려면 사용할 수 있는 함수는?

정답 및 해설 : sample 또는 sample.int 함수

2. 자유도가 5인 카이제곱분포에서 100개의 난수를 발생하여 히스토그램을 그려보아라. 오른쪽으로 긴 꼬리를 가진(오른쪽으로 치우친; skewed to the right)인 분포 모양이 확인되는가?

정답 및 해설 : hist(rchisq(100,5))

3. 자유도가 10인 카이제곱분포에서 100개의 난수를 발생하여 평균과 분산을 얻어보아라. 이 값이 이론적인 기댓값과 분산인 10, 20과 비슷한가?

정답 및 해설 :

(아래의 계산 결과는 발생하는 난수에 따라 약간씩 차이가 있음에 유의)

```
> x <- rchisq(100,10)
```

```
> mean(x)
```

```
[1] 10.03819
```

```
> var(x)
```

```
[1] 18.70144
```

4. 일양분포  $U(0,10)$ 에서 난수를 100개 발생하여 평균과 분산을 계산하여라. 이론적인 값과의 차이가 어느 정도 나는지 확인하여라.

정답 및 해설 :

(아래의 계산 결과는 발생하는 난수에 따라 약간씩 차이가 있음에 유의)

```
> x<- runif(100,0,10)
```

```
> mean(x)
```

```
[1] 4.887889
```

```
> var(x)
```

```
[1] 7.941468
```

5.  $(X,Y)$ 가 이변량정규분포를 따른 확률변수라고 하자.  $X$ 와  $Y$ 의 기댓값은 모두 0이고 분산은 둘 다 1이며 상관계수가  $\rho$ 인 경우를 생각하자. 이 경우 이변량정규분포의 난수는 다음과 같이 생성할 수 있다.

1. 표준정규분포에서  $x$ 를 얻는다.
2.  $y$ 를 기댓값  $\rho x$  분산  $1-\rho^2$ 인 정규분포에서 난수를 얻는다.

위의 방법을 사용하여 다양한  $\rho$  값에 따른  $x$ 와  $y$ 의 산점도를 그려보아라.

정답 및 해설 : 자료실의 bi.norm1.r

## ■ 참고자료

- Oscar T.-R. (2013), Introduction to R Studio
- Pruim, R. (2011), Computational Statistics Using R and R Studio
- W.N.Venables, D.M.Smith and the R Core Team(2014), An Introduction to R. Notes on R: A Programming Environment and Data Analysis and Graphics.
- R : <https://www.r-project.org/>
- R Studio : <https://www.rstudio.com/>

## ■ 코스웨어

- R의 기본 코스웨어 학습방법 안내

다음은 R에 대한 기본적인 학습을 할 수 있는 방법을 안내 해 드립니다. 아직 R에 대하여 익숙하지 못한 학습자는 “R의 입문 코스웨어”를 통하여 R에 대하여 익숙해지기 바랍니다.

### 1) [학교 홈페이지에 접속] - [우측하단 “학과” 선택]



## 2) ["정보통계학과" 선택]



## 3) [상단메뉴 학습정보 내의 "학습자료실" 선택]



#### 4) [좌측메뉴의 “코스웨어” 선택]

Seoul National University Information Systems Department website interface. The left sidebar menu has '코스웨어' (Courseware) highlighted with a red box. An arrow points to this menu item. The main content area shows a table of courseware items.

번호	제목	담당자	작성일	조회수
1	컴퓨터공학과 입시자료	박지현	2019-05-29	37
2	수시교과과정 입시자료	박지현	2019-05-27	26
3	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
4	컴퓨터공학과 입시자료	박지현	2019-05-29	37
5	수시교과과정 입시자료	박지현	2019-05-27	26
6	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
7	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
8	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
9	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
10	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
11	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312
12	일반대학과정(일반대학) 입시자료	한종진	2019-05-08	312

#### 5) [좌측메뉴의 “R의 입문 코스웨어” 선택]

Seoul National University Information Systems Department website interface showing the 'R의 입문 코스웨어' (Introduction to R Courseware) page. The left sidebar menu has 'R의 입문 코스웨어' highlighted with a red box. An arrow points to this menu item. The main content area shows a list of courseware items.

번호	제목	담당자	작성일	조회수
1	R의 입문 코스웨어	한종진	2019-05-08	312

6) “R의 입문 코스웨어”는 총10장으로 구성됨



7) 각 장의 좌측메뉴의 학습목차를 참고하여 학습



“R의 입문 코스웨어” 과정을 통하여 R의 기본개념 및 구조를 이해하고 실제 작업에 필요한 자료처리 및 분석기법을 응용할 수 있는 능력을 학습하기 바랍니다.