

## 5.3 정규성검정

### (1) 정규확률그림

```
par(mfrow = c(1,2)) ; n=10                                # n을 늘려가면서
x = rnorm(n, 0, 1)                                           # N(0,1)에서 난수발생
hist(x, probability = T, main = " Normal(0,1)")             # 히스토그램
  curve( dnorm(x) , add = T)                                  # + 그림

qqnorm(x, sub = "Normal")                                     # Q-Q plot
qqline(x)                                                     # + y=x 직선 추가
```

**[실습1]** 모수를 다르게 설정하고, 난수의 개수(n)를 변화시켜 정규분포를 따르는지 확인..

```
par(mfrow = c(2,2)) ; n=10                                # n을 늘려가면서
x = rnorm(n)                                                  # N(0,1)난수발생
  qqnorm(x, sub = "Normal")                                    # Q-Q plot
  qqline(x)

x = runif(n, min = 0, max = 1)                                # U(0,1)난수발생
  qqnorm(x, sub = "Uniform")                                   # Q-Q plot
  qqline(x)

x = rbinom(n, 10, 0.5)                                         # B(10,0.5)
  qqnorm(x, sub = "Binomial")                                  # Q-Q plot
  qqline(x)

x = rpois(n, lambda=100)                                       # P(100)난수발생
  qqnorm(x, sub = "Poisson")                                   # Q-Q plot
  qqline(x)
```

**[실습2]** 모수를 다르게 설정하고, 난수의 개수(n)를 변화시켜 정규분포를 따르는지 확인..

## (2) 히스토그램 (같은 모수에서 같은 개수의 난수를 발생시켜 여러개의 그래프 작성) 근데 난수는 다르게지요.

### ① 정규분포

```
par(mfrow = c(2,3))  
n=10  
for (i in 1 : 6) {  
    y = rnorm(n, mean = 0, sd = 1)      # 여러줄이니 { }  
    hist(y, ylim = c(0,0.5), prob = T)  # N(0, 1)에서 난수 발생  
    curve( dnorm(x, 0, 1), add = T) }    # N(0, 1)에서 실제 분포
```

**[실습3]** 모수를 다르게 설정하고, 난수의 개수(n)를 변화시켜 정규분포를 따르는지 확인..

### ② 일양분포

```
n=10  
for (i in 1 : 6) {  
    y = runif(n, 0, 1)                  # U(0,1)에서 난수 발생, 평균 = 0.5, 분산=1/12  
    hist(y, ylim = c(0,0.5), prob = T)  # 상황에 맞게 y의 범위 조정  
    curve( dnorm( x, 0.5, 0.29), add = T) }
```

**[실습4]** 모수를 다르게 설정하고, 난수의 개수를 변화시켜 정규분포를 따르는지 확인..

### ③ 이항분포

```
n=10  
for (i in 1 : 6) {  
    y = rbinom(n, 10, 0.5)              # B(10, 0.5)에서 난수 발생, 평균 = 5, 분산 = 2.5  
    hist(y, ylim = c(0,0.5), prob = T)  
    curve( dnorm( x, 5, 1.58), add = T) }
```

**[실습5]** 모수를 다르게 설정하고, 난수의 개수를 변화시켜 정규분포를 따르는지 확인..

### ④ 포아송분포

```
n=10  
for (i in 1 : 6) {  
    y = rpois(n, 4)                     # Poisson(4)에서 난수 발생, # 평균=4, 분산=4  
    hist(y, ylim = c(0,0.5), prob = T)  
    curve( dnorm( x, 4, 2), add = T) }   # N(4, 2^2)에서 실제 분포
```

**[실습6]** 모수를 다르게 설정하고, 난수의 개수를 변화시켜 정규분포를 따르는지 확인..

⑤ 지수분포 ( $f(x) = \lambda e^{-\lambda x}, 0 < x < \infty$ , 평균 =  $\frac{1}{\lambda}$ , 분산 =  $\frac{1}{\lambda^2}$ )

```
dexp (1 , 1)
pexp (1 , 1, lower.tail = T)
pexp (2, 1) - pexp (1, 1)          #  $P(1 \leq x \leq 2)$ 
```

**[실습7]** 모수를 변화시켜 여러 개의 확률계산을 해보시기..

```
x = seq(0,20, by=0.01)
y1 = dexp (x , 1)
plot(x, y1, type = "l", col = "red")

y2 = dexp (x , 5)
y3 = dexp (x , 10)
plot(x, y1, type = "l", col = "2", ylim=c(0,0.8))
  lines(x, y2, type = "l", col = "4")
  lines(x, y3, type = "l", col = "6")
```

**[실습8]** 모수를 변화시켜 여러 가지를 그려보고 1개만 저장.....(다양한 옵션을 사용)

```
n=10
for (i in 1 : 6) {
  y = rexp(n, 1)          # exp(1)에서 난수 발생, # 평균=1, 분산=1
  hist(y, ylim = c(0,0.5), prob = T)
  curve( dnorm( x, 1, 1), add = T ) }    # N(1, 1^2)에서 실제 분포
```

**[실습9]** 모수를 다르게 설정하고, 난수의 개수를 변화시켜 정규분포를 따르는지 확인...

```
rx = rexp(100, 5)      # Exp(5)에서 100개의 난수발생
mean(rx) ; sd(rx)
```

**[실습10]** 모수를 다르게 설정하고, 난수의 개수를 변화시켜 평균과 분산이 참값으로 다가가는지 확인..

```
hist(rx , probability = T, xlim=c(0,3), ylim=c(0, 4), main = " Exponential(5)")
  curve( dexp(x, 5), add = T)
```

**[실습11]** 모수를 다르게 설정하고, 난수의 개수를 변화시켜 그래프의 형태가 참값으로 다가가는지 확인..

### (3) Kolmogorov-Smirnov test (아마 비모수통계학에서 배울 듯)

n=10

y = rnorm(n, mean = 0, sd = 1) # N(0, 1)에서 난수 발생

ks.test(y, "pnorm", mean = 0, sd = 1)

### [과제13] ( [실습1] ~ [실습11] 실습내용 )

첨부파일 : 학번이름13.hwp (예 : 20192260홍길동13.hwp)

- R console 창에서의 프로그램
- 그래프 창의 그래프의 변화 해석