R 이론적 확률분포 1

1. 확률함수

- 단위구간 내에서 어떤 사건이 평균 μ 회 발생하고, 확률변수 X를 사건의 발생횟수라고 할 때

  X ~ Poisson (μ)
- 확률질량함수는 다음과 같음

$$P_r(x=k)=\frac{\mu^k}{k!}e^{-\mu},\quad 0, 1, 2, \cdots$$

- 포아송분포의 모수는 평균(μ)
- 포아송분포의 모양은 평균이 작을 때는 좌우비대칭이나 평균이 증가함에 따라 평균을 중심으로 좌우대칭의 모양으로 변함. 즉, 정규분포에 가까워지고, 분산은 커짐

b1-ch4-5-rev.R

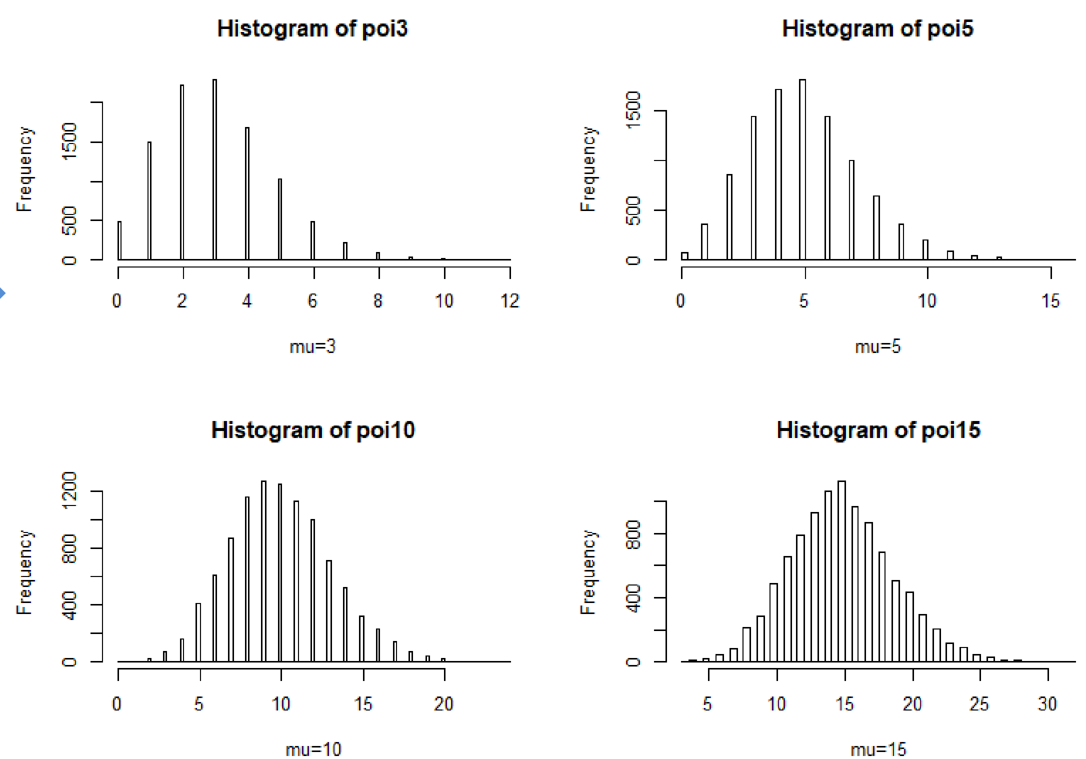
```
set.seed(12345)

n<-10000;

poi3<-rpois(n, 3)
poi5<-rpois(n, 5)
poi10<-rpois(n, 10)
poi15<-rpois(n, 15)

par(mfrow=c(2,2))

hist(poi3, breaks=100, xlab="mu=3")
hist(poi5, breaks=100, xlab="mu=5")
hist(poi10, breaks=100, xlab="mu=10")
hist(poi15, breaks=100, xlab="mu=15")
```

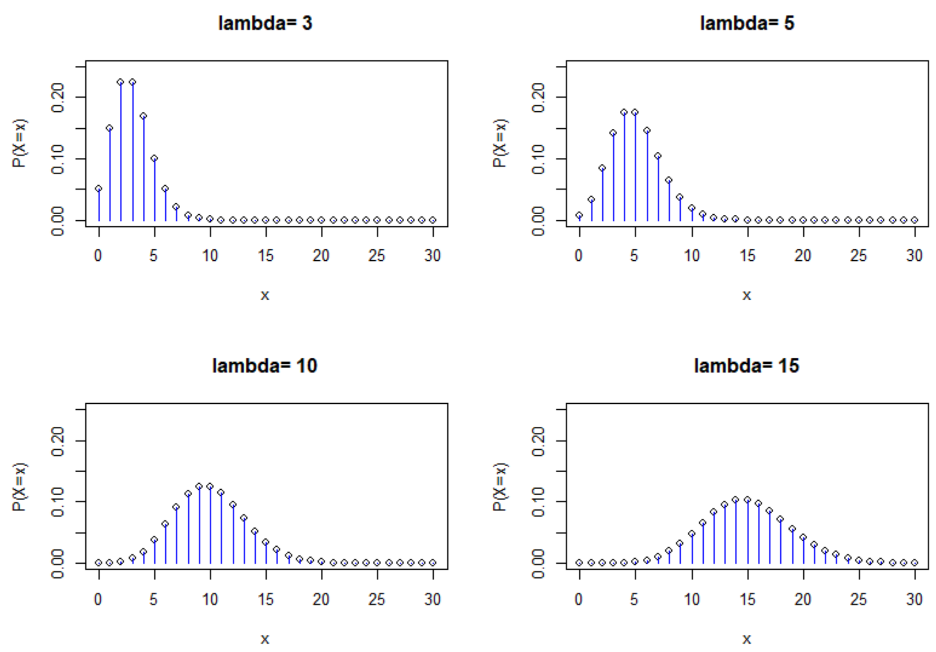


b1-ch4-6.R

```
par(mfrow=c(2,2))

lambda_list<-c(3, 5, 10, 15) # 평균
x_list<-30 # 발생횟수를 x축에 보여줄 최대값

for (i in 1:length(lambda_list)) {
   p_x<-dpois(x=0:x_list,lambda_list[i])
   plot(x=0:x_list, p_x, xlab="x", ylab="P(X=x)", ylim=c(0, 0.25),
       xlim=c(0,x_list), main=paste("lambda=", lambda_list[i]))
   x_seq<-seq(0,x_list,1)
   lines(x_seq, p_x, type="h", col="blue")
}
```



2. 확률분포표

c \ μ	0.02	0.04	0.06	0.08	0.10	0.20	0.30	0.40
0	0.980	0.961	0.942	0.923	0.905	0.819	0.741	0.670
1	1.000	0.999	0.998	0.997	0.995	0.982	0.963	0.938
2	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.992
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

b1-ch4-7.R

```
poi11<-rep(NA,5);poi12<-rep(NA,5);poi13<-rep(NA,5);poi14<-rep(NA,5)
poi15<-rep(NA,5);poi16<-rep(NA,5);poi17<-rep(NA,5);poi18<-rep(NA,5)
poi11[1]<-ppois(0, 0.02);poi12[1]<-ppois(0, 0.04)
poi13[1]<-ppois(0, 0.06);poi14[1]<-ppois(0, 0.08)
poi15[1]<-ppois(0, 0.1);poi16[1]<-ppois(0, 0.2)
poi17[1]<-ppois(0, 0.3);poi18[1]<-ppois(0, 0.4)
for(i in 2:5) {  poi11[i]<-ppois(i-1, 0.02) }
for(i in 2:5) {  poi12[i]<-ppois(i-1, 0.04) }
for(i in 2:5) {  poi13[i]<-ppois(i-1, 0.06) }
for(i in 2:5) {  poi14[i]<-ppois(i-1, 0.08) }
for(i in 2:5) {  poi15[i]<-ppois(i-1, 0.1) }
for(i in 2:5) {  poi16[i]<-ppois(i-1, 0.2) }
for(i in 2:5) {  poi17[i]<-ppois(i-1, 0.3) }
for(i in 2:5) {  poi18[i]<-ppois(i-1, 0.4) }
round((poi<-cbind(poi11,poi12,poi13,poi14,poi15,poi16,poi17,poi18)),digits=3)
```

```
> round((poi<-cbind(poi11,poi12,poi13,poi14,poi15,poi16,poi17,poi18)),digits=3)
     poi11 poi12 poi13 poi14 poi15 poi16 poi17 poi18
[1,]  0.98 0.961 0.942 0.923 0.905 0.819 0.741 0.670
[2,]  1.00 0.999 0.998 0.997 0.995 0.982 0.963 0.938
[3,]  1.00 1.000 1.000 1.000 1.000 0.999 0.996 0.992
[4,]  1.00 1.000 1.000 1.000 1.000 1.000 1.000 0.999
[5,]  1.00 1.000 1.000 1.000 1.000 1.000 1.000 1.000
```

1. 확률함수

- 성공의 확률이 p이고 실패의 확률이 q(q = 1–p)인 베르누이 시행을 독립적으로 n번 반복하였을 때 나타나는 성공의 횟수를 확률변수 X 라고 할 때

  X~B(n,p)

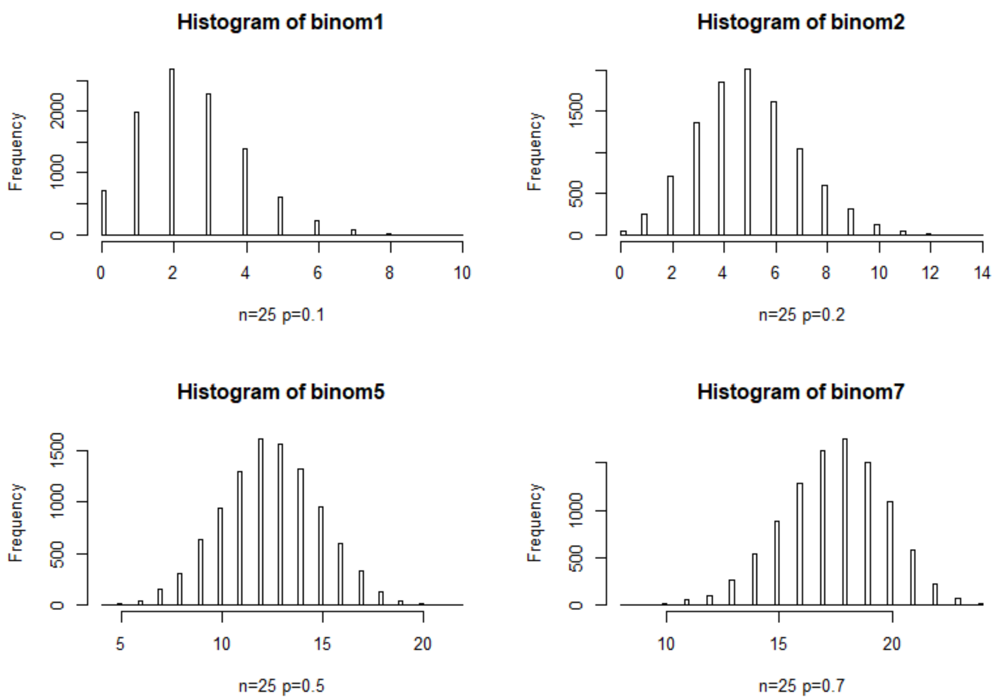
- 확률질량함수는 다음과 같음

$$P_r(X = k) = \binom{n}{k} P^k q^{n-k}, \; k = 0, 1, 2, \cdots, n$$

- 이항분포의 모수는 각 시행에서의 성공확률 p와 시행횟수 n
- 이항분포는 n과 p의 크기에 따라 모양이 결정되는데 p가 0.5이면 평균($\mu = np = \frac{n}{2}$)을 중심으로 좌우 대칭인 분포를 가지며, p=0.5이고 n이 커질 때 좌우대칭의 정규분포와 유사함

b1-ch4-2-rev.R

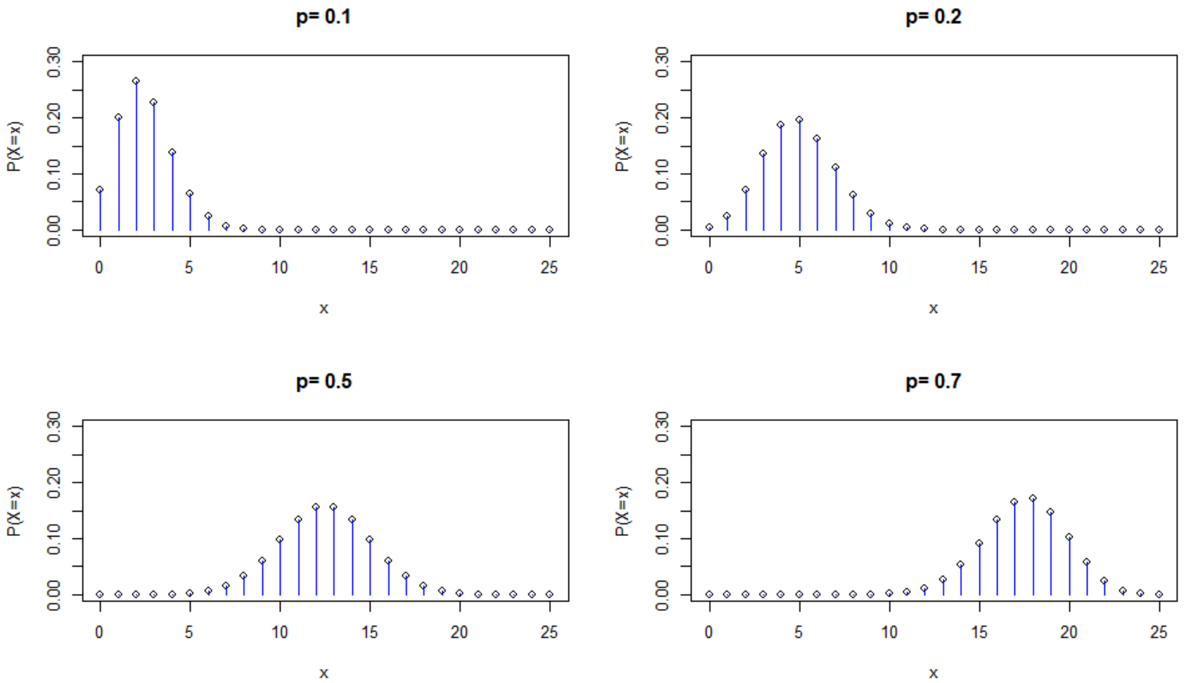
```
set.seed(12345)
r<-10000
binom1<-rbinom(r, 25, 0.1)
binom2<-rbinom(r, 25, 0.2)
binom5<-rbinom(r, 25, 0.5)
binom7<-rbinom(r, 25, 0.7)
par(mfrow=c(2,2))
hist(binom1, breaks=100, xlab="n=25 p=0.1")
hist(binom2, breaks=100, xlab="n=25 p=0.2")
hist(binom5, breaks=100, xlab="n=25 p=0.5")
hist(binom7, breaks=100, xlab="n=25 p=0.7")
```



b1-ch4-3.R

```
par(mfrow=c(2,2))
n<-25 # 시행횟수
p_list<-c(0.1, 0.2, 0.5, 0.7) # 발생확률

for (i in 1:length(p_list)) {
   p_x<-dbinom(x=0:n, n, p_list[i])
   plot(x=0:n, p_x, xlab="x", ylab="P(X=x)",
       ylim=c(0, 0.3), xlim=c(0,n), main=paste("p=", p_list[i]))
   x_seq<-seq(0,n,1)
   lines(x_seq, p_x, type="h", col="blue")
}
```



2. 확률분포표

n=5

a	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
0	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000
1	.919	.737	.528	.337	.187	.087	.031	.007	.000	.000
2	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001
3	1.00	.993	.969	.913	.813	.663	.472	.263	.081	.023
4	1.00	1.00	.998	.990	.969	.922	.832	.672	.410	.226

b1-ch4-4.R

```
binom11<-rep(NA,5);binom12<-rep(NA,5);binom13<-rep(NA,5);binom14<-rep(NA,5)
binom15<-rep(NA,5);binom16<-rep(NA,5);binom17<-rep(NA,5);binom18<-rep(NA,5);binom19<-rep(NA,5)
binom11[1]<-pbinom(0, 5, 0.1);binom12[1]<-pbinom(0, 5, 0.2);binom13[1]<-pbinom(0, 5, 0.3);binom14[1]<-pbinom(0, 5, 0.4)
binom15[1]<-pbinom(0, 5, 0.5);binom16[1]<-pbinom(0, 5, 0.6);binom17[1]<-pbinom(0, 5, 0.7);binom18[1]<-pbinom(0, 5, 0.8)
binom19[1]<-pbinom(0, 5, 0.9)
for(i in 2:5) {  binom11[i]<-pbinom(i-1, 5, 0.1) }
for(i in 2:5) {  binom12[i]<-pbinom(i-1, 5, 0.2) }
for(i in 2:5) {  binom13[i]<-pbinom(i-1, 5, 0.3) }
for(i in 2:5) {  binom14[i]<-pbinom(i-1, 5, 0.4) }
for(i in 2:5) {  binom15[i]<-pbinom(i-1, 5, 0.5) }
for(i in 2:5) {  binom16[i]<-pbinom(i-1, 5, 0.6) }
for(i in 2:5) {  binom17[i]<-pbinom(i-1, 5, 0.7) }
for(i in 2:5) {  binom18[i]<-pbinom(i-1, 5, 0.8) }
for(i in 2:5) {  binom19[i]<-pbinom(i-1, 5, 0.9) }
round((binom<-cbind(binom11,binom12, binom13,binom14,binom15, binom16,binom17,binom18, binom19)),digits=3)
```

```
> round((binom<-cbind(binom11,binom12, binom13,binom14,binom15, binom16,binom17,binom18, binom19)),digits=3)
     binom11 binom12 binom13 binom14 binom15 binom16 binom17 binom18 binom19
[1,]   0.590   0.328   0.168   0.078   0.031   0.010   0.002   0.000   0.000
[2,]   0.919   0.737   0.528   0.337   0.187   0.087   0.031   0.007   0.000
[3,]   0.991   0.942   0.837   0.683   0.500   0.317   0.163   0.058   0.009
[4,]   1.000   0.993   0.969   0.913   0.812   0.663   0.472   0.263   0.081
[5,]   1.000   1.000   0.998   0.990   0.969   0.922   0.832   0.672   0.410
```