

生物資源解析学演習 2022 Lecture 2

R の使い方基礎 (2): 確率分布 + モンテカルロ法

北門 利英 (東京海洋大学海洋生物資源学科)

2022 年 10 月 11 日

Contents

1	確率分布	1
1.1	R における確率分布に関連した表記	1
2	2 項分布	2
2.1	確率分布 (確率関数) と累積確率分布関数のグラフ化	2
2.2	2 項分布に従う例題: 哺乳類の雌の個体数変動	3
2.3	HW1 上記の設定で以下の考察のポイントについて検討しなさい	6
3	ポアソン分布	6
3.1	確率分布 (確率関数) と累積確率分布関数のグラフ化	6
3.2	HW2 雌個体数動態の発展	7
3.3	HW3 雌個体数動態の発展	7
4	一様分布	9
4.1	一様乱数を用いたモンテカルロ法による円周率の計算	9
4.2	HW4 繰り返し数を 10^5 までとするとどうでしょうか?	12
5	個体群動態モデル (参考までに, また今後扱います)	13
5.1	Pella-Tomlinson 余剰生産モデル (確率変動なし)	13
5.2	Pella-Tomlinson 余剰生産モデル (確率変動あり)	15

Attention:

- 授業に関してわからないことがあれば, 授業中あるいは授業後に遠慮なく質問してください

Point:

- Lecture 2 では確率分布およびモンテカルロ法について学びます.

1 確率分布

1.1 R における確率分布に関連した表記

R には統計解析やモデリングで利用する種々の確率分布に対応した関数が用意されている. これらは, 各種の確率分布に従う乱数の発生その他, 確率関数 (確率密度関数), 累積確率分布, そしてパーセント点の評価に分類される. 例えば, 2 項分布 $Bin(N, p)$ を例にとると以下の通り. これと同様にして, 頭文字を r , d , p , q とすることで他の確率分布にも適用できる.

```
rbinom(n=乱数の数, size=N, prob=p) #独立な乱数の生成
```

```
dbinom(x=実現値の値, size, prob, log = FALSE) #確率関数の値
```

```
pbinom(q, size, prob) #累積確率の値
qbinom(p, size, prob) #パーセント点の値
```

Table 1: それぞれの確率分布に対応した関数

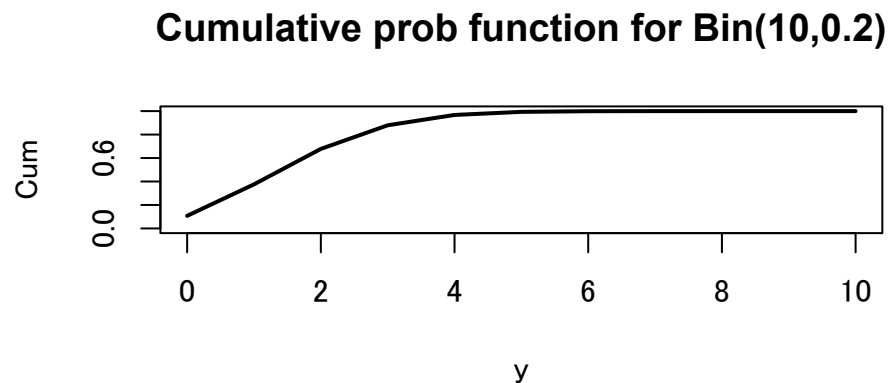
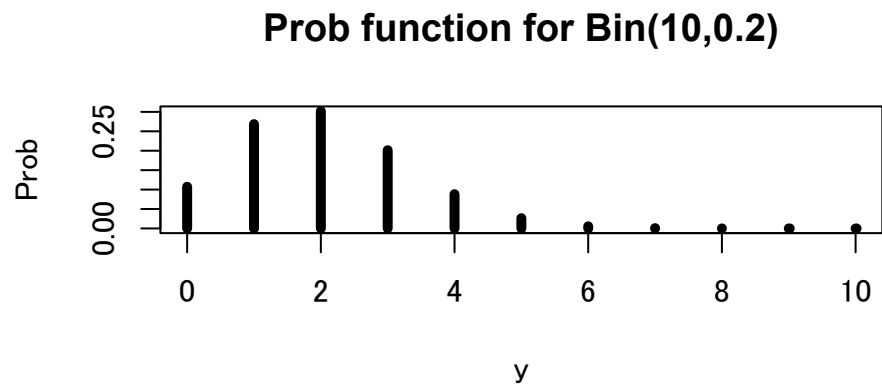
離散型確率分布	
<code>_binom(, N, p)</code>	2 項分布 $Bin(N, p)$
<code>_pois(, lambda)</code>	ポアソン分布 $Pois(\lambda)$
<code>_nbinom(, N, p)</code>	負の 2 項分布 $NB(N, p)$
<code>_geom(, p)</code>	幾何分布 $Geom(p)$
<code>_hyper(, m=M, n=N-M, k=n)</code>	超幾何分布 $HG(N, M, n)$
<code>_multinom(, N, p)</code>	多項分布 $Multinom(N, p)$
連続型確率分布	
<code>_unif(, min=a, max=b)</code>	一様分布 $U(a, b)$
<code>_norm(, mu, sigma)</code>	正規分布 $N(\mu, \sigma^2)$
<code>_lnorm(, mu, sigma)</code>	対数正規分布 $LN(\mu, \sigma^2)$
<code>_gamma(, shape=a, scale=b)</code>	ガンマ分布 $Ga(a, b)$ with $E[Y] = ab$
<code>_exp(, scale=sigma)</code>	ガンマ分布 $Exp(\sigma)$ with $E[Y] = \sigma$
<code>_chisq(, df=n)</code>	カイ 2 乗分布 $\chi^2(n)$
<code>_beta(, shape1=a, shape2=b)</code>	ベータ分布 $\beta(n_1, n_2)$
<code>_t(, df=n)</code>	t 分布 $t(n)$
<code>_f(, df1=n1, df2=n2)</code>	t 分布 $F(n_1, n_2)$

2 2 項分布

2.1 確率分布 (確率関数) と累積確率分布関数のグラフ化

- $Y \sim Bin(10, 0.2)$ に対して $Pr(Y = y)$ および $P(Y \leq y) (y = 0, 1, \dots, 10)$ の計算

```
N <- 10
p <- 0.2
Yrange <- 0:10
Prob <- dbinom(Yrange, N, p)
Cum <- pbinom(Yrange, N, p)
par(mfrow=c(2,1))
plot(Yrange, Prob, type="h", lwd=5, xlab="y",
     main="Prob function for Bin(10,0.2)")
plot(Yrange, Cum, type="l", lwd=2, xlab="y",
     main="Cumulative prob function for Bin(10,0.2)", ylim=c(0,1))
```



- $Y \sim \text{Bin}(20, 0.1)$ に対して $Pr(Y = y)$ および $P(Y \leq y) (y = 0, 1, \dots, 20)$ の計算

Do by yourself!

```
plot(Yrange, Prob, type="h", lwd=5, xlab="y",
     main=paste("2 項分布 Bin(", N, ", ", p, ") の確率関数"))
```

2.2 2 項分布に従う例題：哺乳類の雌の個体数変動

Notation

- N_t : t 年に初めの雌の数
- S_t : t 年における生き残り雌の数
- P_t : t 年の生き残り雌のうち、妊娠雌の数 (= 出産数)
- B_t : 出産数 P_t 個のうち雌の数
- 1 年あたりの死亡率 d ($=0.2$)
- 1 年あたりの妊娠率 b ($=0.4$) [雌 1 個体から 1 個体出産]

Stochastic nature

$$\begin{aligned}
 S_t | N_t &\sim \text{Bin}(N_t, 1-d) \\
 P_t | S_t &\sim \text{Bin}(S_t, b) \\
 B_t | P_t &\sim \text{Bin}(P_t, 0.5) \\
 N_{t+1} &= S_t + B_t
 \end{aligned}$$

- 1年目の雌個体数 Nf (=5)
- 20年予測 (TT=10)
- シミュレーション繰り返し数 Nsim (=100)

#パラメータ値の設定

```

d <- 0.2
b <- 0.4
TT <- 10
TT1 <- TT-1
Nsim <- 100
Nf <- array(0, c(Nsim, TT))
Nf[,1] <- 5

```

#1回だけの計算

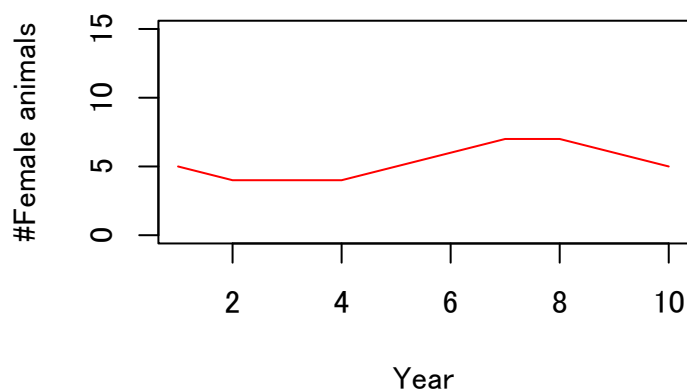
```

for(t in 1:TT1) {
  Surv <- rbinom(1, Nf[1, t], 1-d)
  Preg <- rbinom(1, Surv, b)
  Birthf <- rbinom(1, Preg, 0.5)
  Nf[1, t+1] <- Surv + Birthf
}
Nf[1,]

```

```
[1] 5 4 4 4 5 6 7 7 6 5
```

```
plot(Nf[1,], type="l", col="red", xlab="Year", ylab="#Female animals", ylim=c(0,15))
```

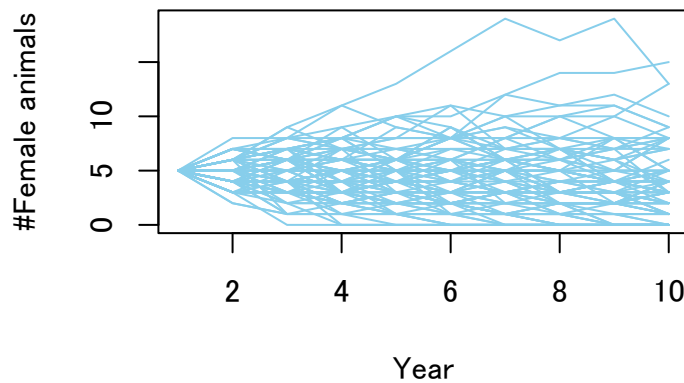


計算の繰り返し

```

for(i in 1:Nsim){
  for(t in 1:TT1){
    Surv <- rbinom(1, Nf[i, t], 1-d)
    Preg <- rbinom(1, Surv, b)
    Birthf <- rbinom(1, Preg, 0.5)
    Nf[i, t+1] <- Surv + Birthf
  }
}
plot(Nf[1,], type="l", col="skyblue", ylim=c(0,max(Nf)),
      xlab="Year", ylab="#Female animals", main=paste("TT=",TT))
for(i in 1:Nsim){
  points(Nf[i,], type="l", col="skyblue")
}

```

TT= 10

Nf[,TT]

```

[1] 0 2 8 0 2 5 4 4 3 4 10 1 0 8 8 0 3 3 3 1 0 3 0 5 5
[26] 1 3 0 0 8 3 0 2 4 0 1 3 7 3 1 9 13 2 9 3 8 0 4 2 3
[51] 0 3 2 5 2 5 1 3 3 5 4 0 1 0 4 2 4 15 2 4 8 0 1 7 5
[76] 1 5 7 0 0 13 8 1 4 3 2 7 0 5 5
[ reached getOption("max.print") -- omitted 10 entries ]

```

mean(Nf[,TT])

[1] 3.54

Nf[,TT]==0

```

[1] TRUE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] TRUE FALSE FALSE TRUE FALSE FALSE FALSE FALSE TRUE FALSE TRUE FALSE
[25] FALSE FALSE FALSE TRUE TRUE FALSE FALSE TRUE FALSE FALSE TRUE FALSE
[37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE

```

```
[49] FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[61] FALSE TRUE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE
[73] FALSE FALSE FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE
[85] FALSE FALSE FALSE TRUE FALSE FALSE
[ reached getOption("max.print") -- omitted 10 entries ]
```

```
mean(Nf[,TT]==0)
```

```
[1] 0.2
```

10 年後に絶滅する確率は？

```
paste("Extinction prob = ", mean(Nf[,TT] == 0))
```

```
[1] "Extinction prob = 0.2"
```

2.3 HW1 上記の設定で以下の考察のポイントについて検討しなさい

- 年間死亡率が変わると絶滅確率も変化するでしょう。その関係をグラフで表現してみてください
- 年間妊娠率が変わると絶滅確率も変化するでしょう。その関係をグラフで表現してみてください
- 個体群は、いずれは絶滅するのでしょうか？例えば TT=20,50,100,1000 などと設定し、考察してみてください

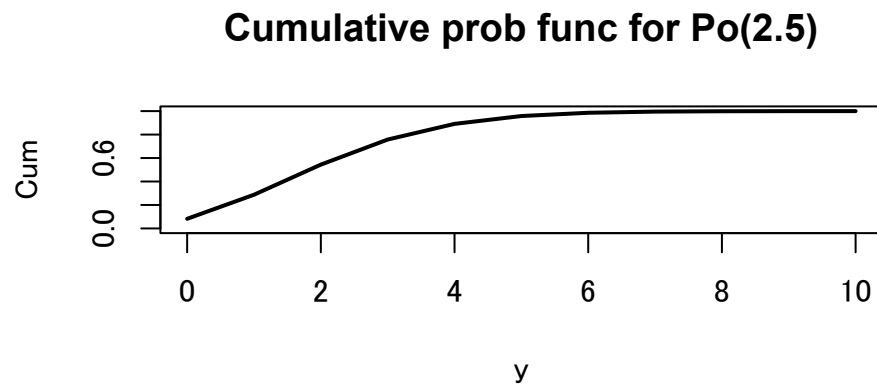
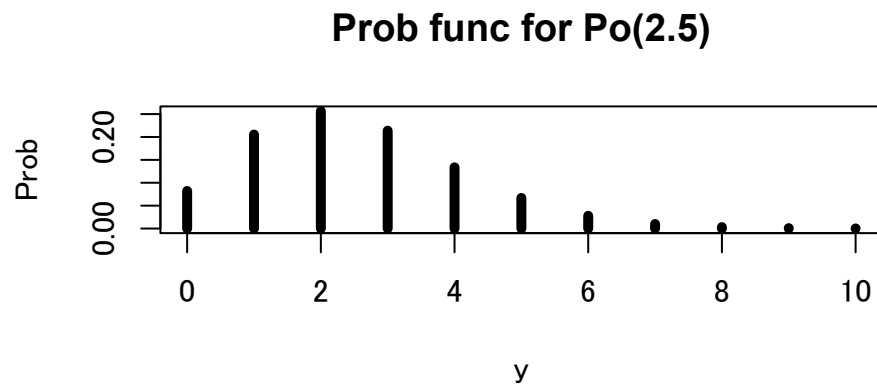
3 ポアソン分布

```
rpois(n=乱数の数, lambda) #独立な乱数の生成
dpois(x=実現値の値, lambda, log = FALSE) #確率関数の値
ppois(q, lambda=lambda) #累積確率の値
qpois(p, lambda=lambda) #パーセント点の値
```

3.1 確率分布 (確率関数) と累積確率分布関数のグラフ化

$Y \sim Po(2.5)$ に対して $Pr(Y = y)$ および $P(Y \leq y) (y = 0, 1, \dots, 10)$ の計算

```
lambda <- 2.5
Yrange <- 0:10
Prob <- dpois(Yrange, lambda)
Cum <- ppois(Yrange, lambda)
par(mfrow=c(2,1))
plot(Yrange, Prob, type="h", lwd=5, xlab="y",
     main="Prob func for Po(2.5)")
plot(Yrange, Cum, type="l", lwd=2, xlab="y",
     main="Cumulative prob func for Po(2.5)", ylim=c(0,1))
```



$Y \sim Po(5)$ に対して $Pr(Y = y)$ および $P(Y \leq y)(y = 0, 1, \dots, 20)$ の計算

Do by yourself!

3.2 HW2 雌個体数動態の発展

$Y \sim Po(2.5)$ と $Y \sim Bin(50, 0.05)$ の比較

Do by yourself!

3.3 HW3 雌個体数動態の発展

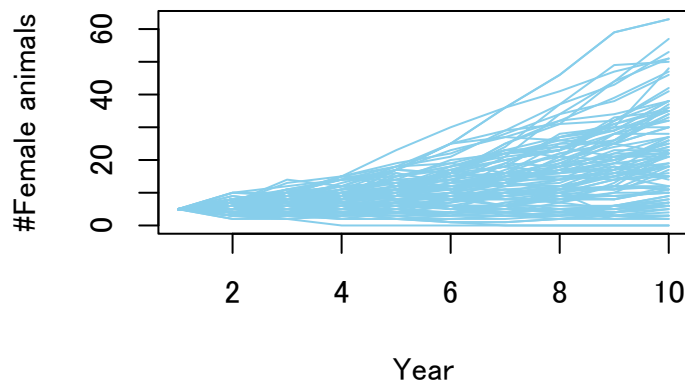
- 上記の雌個体数の動態では雌 1 個体から 1 個体出産と仮定した
- 次に雌 1 個体から複数個体が生まれる可能性を考慮したい
- 出生数が $1+Po(0.2)$ に従うときに、雌個体数の計算はどのように行えばよいか
- またこの場合、絶滅確率に変化はあるか? (T=20, 50 くらいで計算)

```

d <- 0.2
b <- 0.4
lambda <- 0.2
TT <- 10; TT1<-TT-1
Nsim <- 100
Nf <- array(0, c(Nsim, TT))
Nf[,1] <- 5
for(i in 1:Nsim){
  for(t in 1:TT1){
    Surv <- rbinom(1, Nf[i, t], 1-d)
    Preg <- rbinom(1, Surv, b)
    if(Preg==0){Birthf <- 0} else {Birthf <- Preg + sum(rpois(Preg, lambda))}
    Nf[i, t+1] <- Surv + Birthf
  }
}
plot(Nf[1,], type="l", col="skyblue", xlab="Year", ylab="#Female animals",
      ylim=c(0,max(Nf)), main=paste("TT=",TT))
for(i in 1:Nsim){
  points(Nf[i,], type="l", col="skyblue")
}

```

TT= 10



```
paste("Extinction prob = ", mean(Nf[,TT] == 0))
```

```
[1] "Extinction prob = 0.02"
```

```

TT <- 20; TT1<-TT-1
Nf <- array(0, c(Nsim, TT))
Nf[,1] <- 5
for(i in 1:Nsim){
  for(t in 1:TT1){
    Surv <- rbinom(1, Nf[i, t], 1-d)

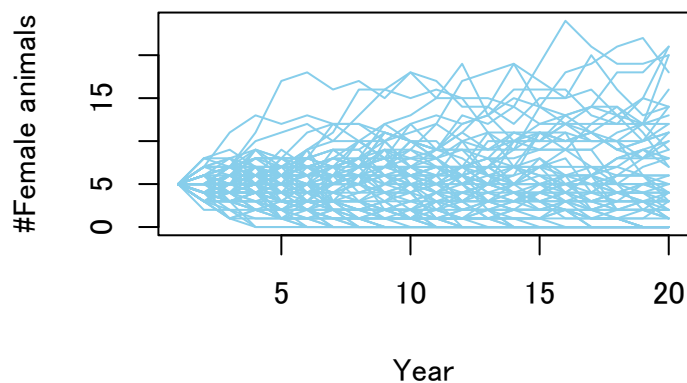
```

```

Preg <- rbinom(1, Surv, b)
if(Preg==0){Birthf <- 0} else {
  Birth <- Preg + sum(rpois(Preg, lambda))
  Birthf <- rbinom(1,Birth,0.5)
}
Nf[i, t+1] <- Surv + Birthf
}
}
plot(Nf[1,], type="l", col="skyblue", xlab="Year", ylab="#Female animals",
      ylim=c(0,max(Nf)), main=paste("TT=",TT))
for(i in 1:Nsim){
  points(Nf[i,], type="l", col="skyblue")
}

```

TT= 20



```
paste("Extinction prob = ", mean(Nf[,TT] == 0))
```

```
[1] "Extinction prob = 0.43"
```

4 一様分布

$$Y \sim U(a, b)$$

$$f(y) = \frac{1}{b-a} \quad \text{for } a < y < b$$

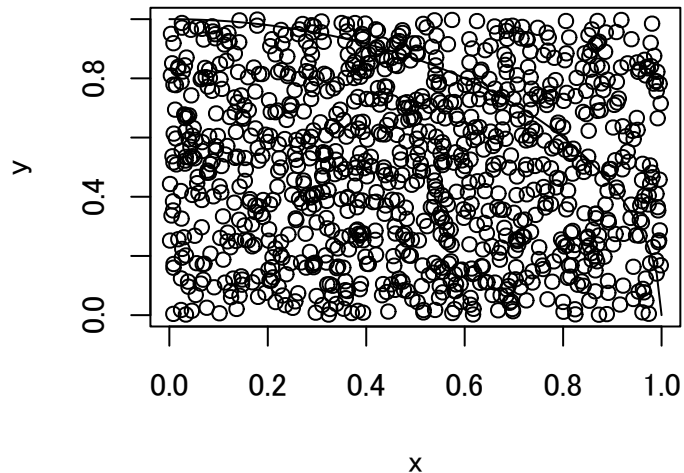
4.1 一様乱数を用いたモンテカルロ法による円周率の計算

$$X \sim U(0, 1)$$

$$Y \sim U(0, 1)$$

$$Pr(X^2 + Y^2 \leq 1) = \pi/4$$

```
Nit <- 1000
x <- runif(Nit, 0, 1)
y <- runif(Nit, 0, 1)
plot(x, y)
curve( sqrt(1-x^2), xlim=c(0,1), add=T)
```



```
4*mean(x^2+y^2 <= 1)
```

```
[1] 3.208
```

関数化すると

```
Calc_pi <- function(Nit, Fig=F)
{
  x <- runif(Nit, 0, 1)
  y <- runif(Nit, 0, 1)
  if(Fig==T){
    plot(x, y, main=paste("#iterations = ", Nit), pch=19, cex=0.3, col="gray")
    curve( sqrt(1-x^2), xlim=c(0,1), add=T, col="red")
  }
  4*mean(x^2+y^2 <= 1)
}
```

```
par(mfrow=c(2,2))
Calc_pi(100, Fig=T)
```

```
[1] 2.96
```

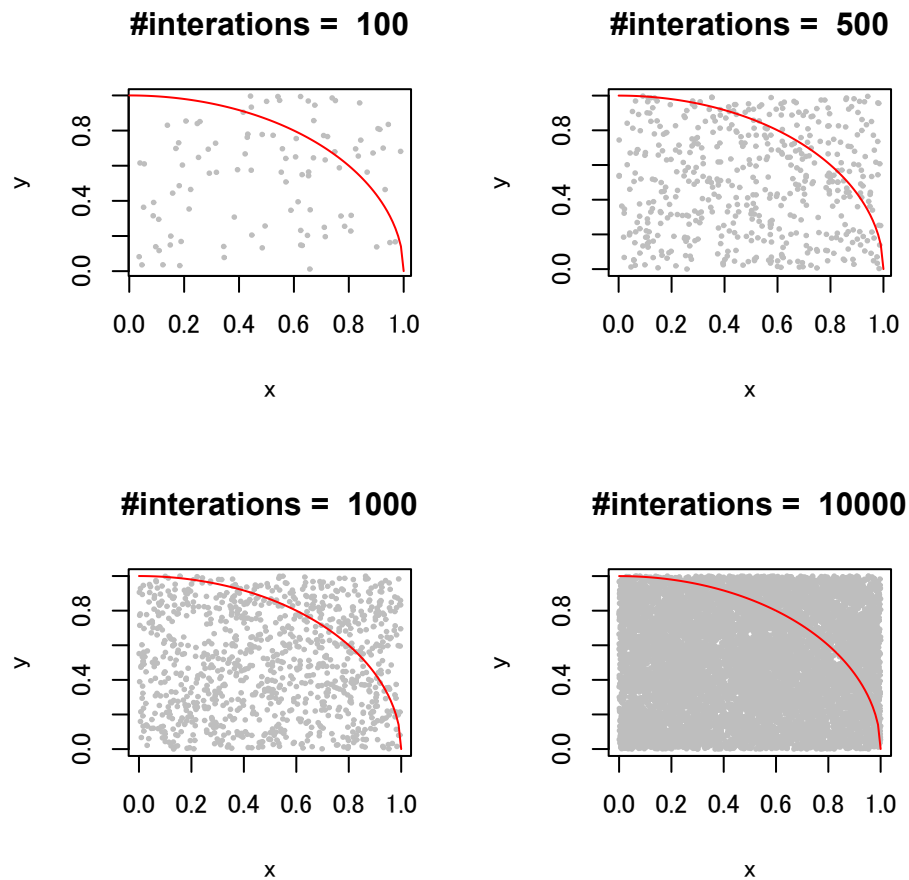
```
Calc_pi(500, Fig=T)
```

```
[1] 3.2
```

```
Calc_pi(1000, Fig=T)
```

```
[1] 3.136
```

```
Calc_pi(10000, Fig=T)
```



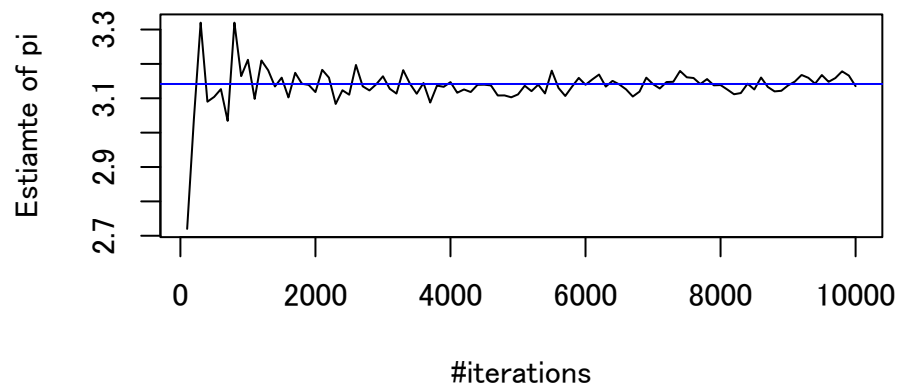
```
[1] 3.144
```

```
Nit_vec <- seq(100,10000,100)
```

```
Pi_vec <- sapply(Nit_vec, Calc_pi)
```

```
plot(Nit_vec, Pi_vec, type="l", xlab="#iterations", ylab="Estiamte of pi")
```

```
abline(pi, 0, col="blue")
```

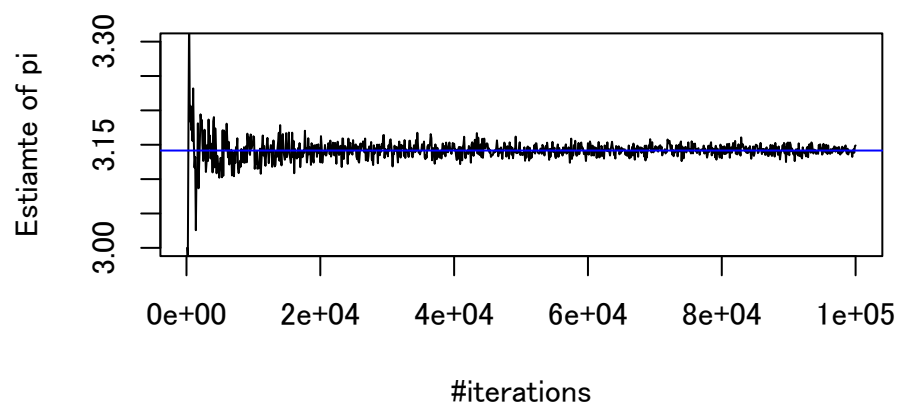


4.2 HW4 繰り返し数を 10^5 までとするとどうでしょうか？

```
Nit_vec <- seq(100,100000,100)
Pi_vec <- sapply(Nit_vec, Calc_pi)

plot(Nit_vec, Pi_vec, type="l", xlab="#iterations", ylab="Estiamte of pi",
     main="Nit is up to 10^5", ylim=c(3,3.3))
abline(pi, 0, col="blue")
```

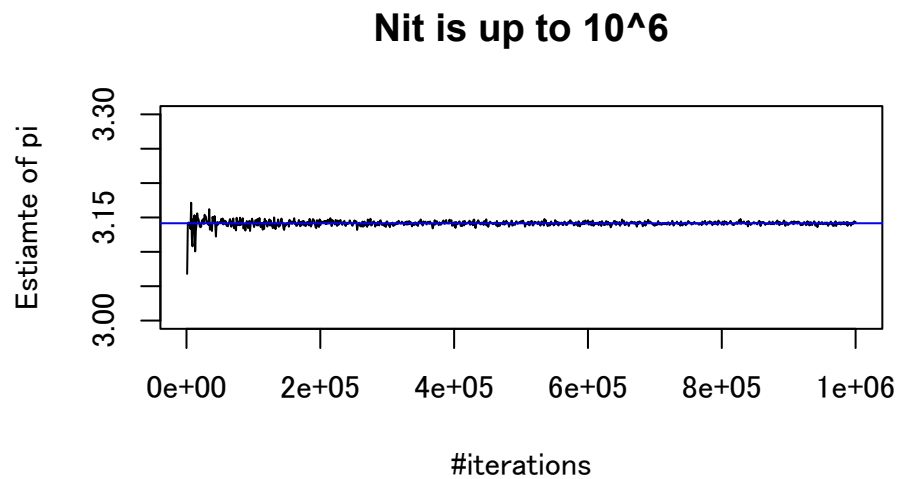
Nit is up to 10^5



```
Nit_vec <- seq(1000,1000000,1000)
Pi_vec <- sapply(Nit_vec, Calc_pi)

plot(Nit_vec, Pi_vec, type="l", xlab="#iterations", ylab="Estiamte of pi",
```

```
main="Nit is up to 10^6", ylim=c(3,3.3))
abline(pi, 0, col="blue")
```



5 個体群動態モデル (参考までに, また今後扱います)

5.1 Pella-Tomlinson 余剰生産モデル (確率変動なし)

5.1.1 モデルの定義

$$P_{t+1} = P_t + rP_t \left\{ 1 - \left(\frac{P_t}{K} \right)^z \right\} - C_t$$

5.1.2 初歩的に漸化式を利用して個体群動態を計算

```
r <- 0.2
K <- 10000
z <- 1
TT <- 50

Catch <- rep(500, TT)

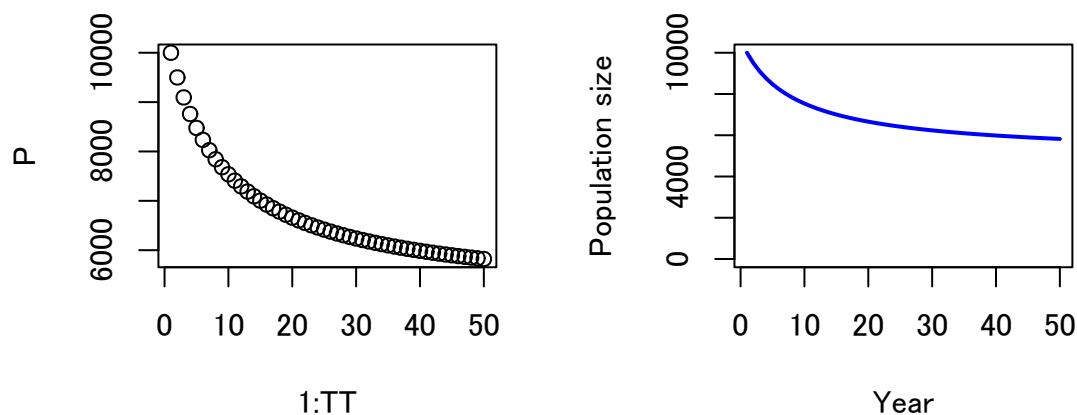
TT1 <- TT-1
P <- numeric(TT)
P[1] <- K
for(t in 1:TT1)
{
  tmp <- P[t] + r*P[t]*(1-(P[t]/K)^z) - Catch[t]
  P[t+1] <- max(tmp, 0.001)
}

print(P, digits=0)
```

```
[1] 10000 9500 9095 8760 8477 8235 8026 7843 7681 7537 7409 7293 7187 7092 7004
[16] 6924 6850 6781 6718 6659 6604 6552 6504 6459 6416 6376 6338 6303 6269 6236
[31] 6206 6177 6149 6123 6097 6073 6050 6028 6007 5987 5967 5949 5931 5913 5897
[46] 5881 5865 5850 5836 5822
```

5.1.3 計算結果の図示

```
par(mfrow=c(1,2))
plot(1:TT, P)
plot(1:TT, P, type="l", lwd=2, col="blue", xlab="Year", ylab="Population size", ylim=c(0,K))
```



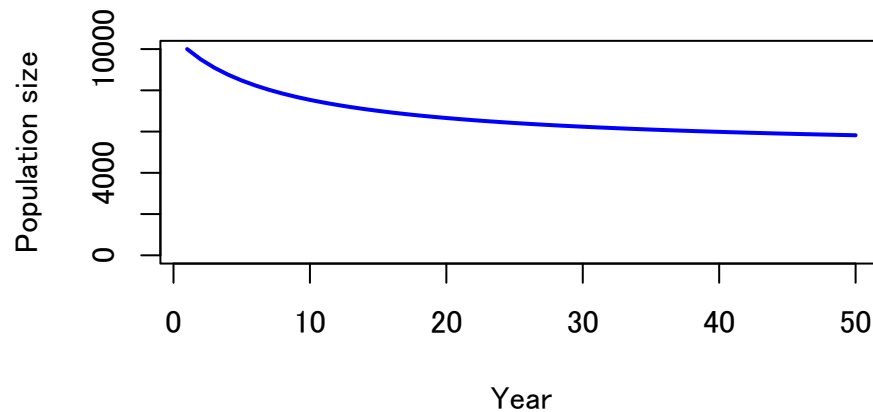
5.1.4 もう少し進んだ方法：個体群動態計算用関数 PDM.PT を作成

```
#Pella-Tomlinson Production model (without stochastic process error)
PDM.PT<-function(r, K, z, P1, TT, Catch)
{
  TT1 <- TT-1
  P <- numeric(TT)
  P[1] <- P1
  for(t in 1:TT1)
  {
    tmp <- P[t] + r*P[t]*(1-(P[t]/K)^z) - Catch[t]
    P[t+1] <- max(tmp, 0.001)
  }
  return(P)
}
```

5.1.5 新しく作った個体群動態計算用関数 PDM.PT を実行

```
Res <- PDM.PT(r=0.2, K=10^4, z=1, P1=10^4, TT=50, Catch=rep(500,TT))
print(Res, digits=0)
```

```
[1] 10000 9500 9095 8760 8477 8235 8026 7843 7681 7537 7409 7293 7187 7092 7004
[16] 6924 6850 6781 6718 6659 6604 6552 6504 6459 6416 6376 6338 6303 6269 6236
[31] 6206 6177 6149 6123 6097 6073 6050 6028 6007 5987 5967 5949 5931 5913 5897
[46] 5881 5865 5850 5836 5822
plot(1:TT, Res, type="l", lwd=2, col="blue", xlab="Year", ylab="Population size", ylim=c(0,K))
```



5.2 Pella-Tomlinson 余剰生産モデル (確率変動あり)

5.2.1 モデルの定義

$$P_{t+1} = \left[P_t + rP_t \left\{ 1 - \left(\frac{P_t}{K} \right)^z \right\} - C_t \right] * \exp(\varepsilon_t), \quad \varepsilon_t \sim N(0, \sigma^2)$$

5.2.2 さっき作った個体群動態計算用関数 PDM.PT を書き換え (上書き更新)

```
#Pella-Tomlinson Production model (with stochastic process error)
PDM.PT<-function(r, K, z, P1, TT, Catch, sigma=0)
{
  TT1 <- TT-1
  P <- numeric(TT)
  Epsilon <- rnorm(n=TT, mean=0, sd=sigma)
  P[1] <- P1
  for(t in 1:TT1)
  {
    tmp <- P[t] + r*P[t]*(1-(P[t]/K)^z) - Catch[t]
    tmp <- tmp*exp(Epsilon[t])
    P[t+1] <- max(tmp, 0.001)
  }
  return(P)
}
```

5.2.3 再定義した PDM.PT を実行

```

K <- 10^4
TT <- 50

Res <- PDM.PT(r=0.2, K=K, z=1, P1=K, TT=50, Catch=rep(500,TT), sigma=0.05)
print(Res, digits=0)

[1]10000 9339 9401 9299 8441 7825 8120 8179 7645 7529 6972 7142 6814 6352 5700
[16] 5823 5840 5326 5596 5628 5792 5646 5951 6029 6080 6283 6117 6231 6238 6230
[31] 5755 5711 4945 5242 5194 5452 5051 5246 5453 5374 5937 5759 5028 4990 4827
[46] 4906 4568 4029 4333 4416

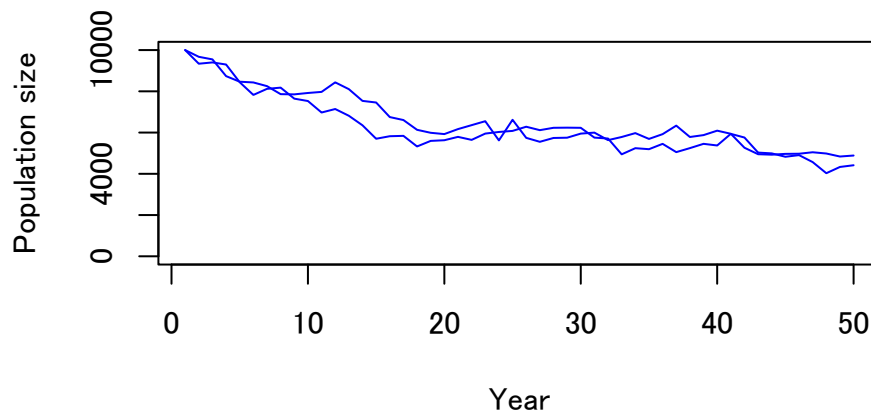
plot(1:TT, Res, type="l", col="blue", xlab="Year", ylab="Population size", ylim=c(0,K))

Res <- PDM.PT(r=0.2, K=K, z=1, P1=K, TT=50, Catch=rep(500,TT), sigma=0.05)
print(Res, digits=0)

[1]10000 9673 9548 8737 8466 8430 8251 7863 7852 7920 7972 8435 8105 7536 7457
[16] 6751 6607 6133 5989 5926 6168 6360 6551 5620 6620 5744 5552 5738 5751 5949
[31] 5998 5633 5785 5973 5692 5923 6336 5785 5877 6092 5941 5264 4950 4927 4966
[46] 4974 5045 4984 4838 4887

points(1:TT, Res, type="l", col="blue")

```



5.2.4 PDM.PT を繰り返し実行して結果を図示

```

Nsim <- 100
Pmat <- array(NA, c(Nsim, TT))
plot(0, type="n", xlab="Year", ylab="Population size", xlim=c(0,TT), ylim=c(0,K))

for(i in 1:Nsim)

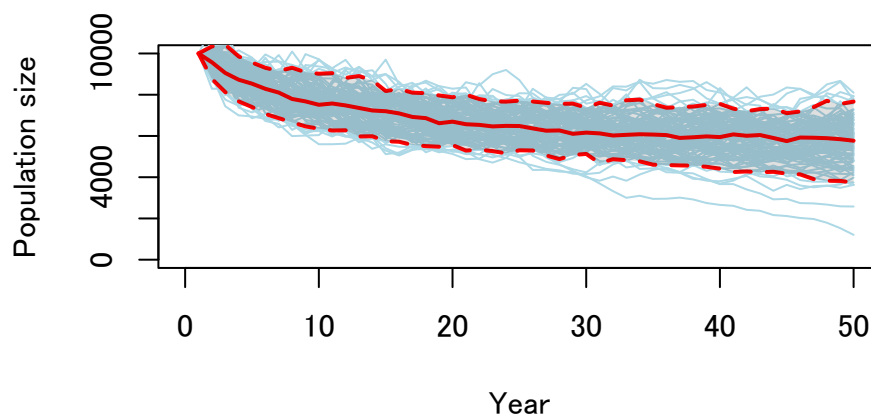
```

```

{
  Pmat[i,] <- PDM.PT(0.2, 10^4, 1, 10^4, 50, rep(500,TT), 0.05)
  points(1:TT, Pmat[i,], type="l", col="lightblue")
}
P.med <- apply(Pmat, 2, median);
P.L5per <- apply(Pmat, 2, quantile, 0.05)
P.U5per <- apply(Pmat, 2, quantile, 0.95)
points(1:TT, P.med, type="l", lwd=2, col="red")
points(1:TT, P.L5per, type="l", lty=2, lwd=2, col="red")
points(1:TT, P.U5per, type="l", lty=2, lwd=2, col="red")

polygon(
  c(1:TT, TT:1), c(P.L5per, rev(P.U5per)),
  col = "#00000020", border = NA)

```



5.2.5 上記を関数化

```

PDM.PT.sim <- function(r, K, z, P1, TT, Catch, sigma=0, Nsim)
{
  Pmat <- array(NA, c(Nsim, TT))
  plot(0, type="n", xlab="Year", ylab="Population size",
       xlim=c(0,TT), ylim=c(0,K), xaxs="i", yaxs="i")

  for(i in 1:Nsim)
  {
    Pmat[i,] <- PDM.PT(0.2, 10^4, 1, 10^4, 50, rep(500,TT), 0.05)
    points(1:TT, Pmat[i,], type="l", col="lightblue")
  }
  P.med <- apply(Pmat, 2, median);
  P.L5per <- apply(Pmat, 2, quantile, 0.05)

```

```

P.U5per <- apply(Pmat, 2, quantile, 0.95)
points(1:TT, P.med, type="l", lwd=2, col="red")
points(1:TT, P.L5per, type="l", lty=2, lwd=2, col="red")
points(1:TT, P.U5per, type="l", lty=2, lwd=2, col="red")

polygon( c(1:TT, TT:1), c(P.L5per, rev(P.U5per)), col = "#00000020", border = NA)

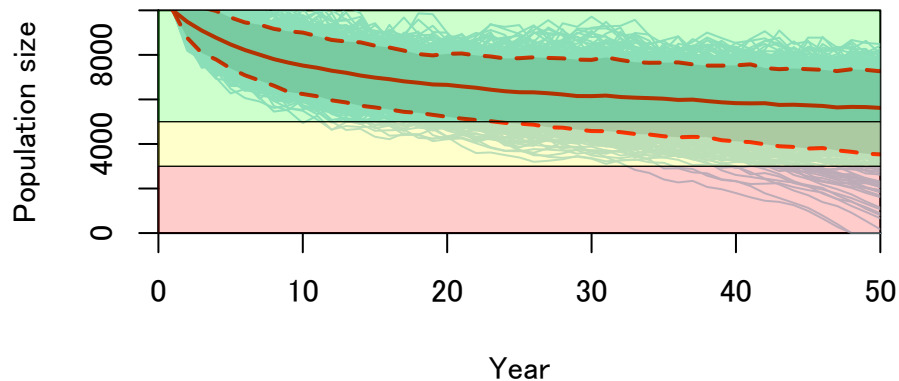
rect(xleft=-1, ybottom=0, xright=TT, ytop=0.3*K, lwd=0, col=rgb(1,0,0,alpha=0.2))
rect(xleft=-1, ybottom=0.3*K, xright=TT, ytop=0.5*K, lwd=0, col=rgb(1,1,0,alpha=0.2))
rect(xleft=-1, ybottom=0.5*K, xright=TT, ytop=K, lwd=0, col=rgb(0,1,0,alpha=0.2))

return(Pmat)
}

```

5.2.6 実行！

```
Res <- PDM.PT.sim(0.2, 10^4, 1, 10^4, 50, rep(500,TT), 0.05, Nsim=1000)
```



```

nf <- layout(matrix(c(1,2,0,0),2,2,byrow=TRUE), c(3,1), c(4,0), FALSE)
par(mar=c(4,4,3,1))
Res <- PDM.PT.sim(0.2, 10^4, 1, 10^4, 50, rep(500,TT), 0.05, Nsim=1000)
par(mar=c(3,0,2,1))
Hist.final <- hist(Res[,TT], breaks=seq(0,1.2*K,500), plot=FALSE)
top <- max(Hist.final$counts)
barplot(Hist.final$counts, axes=FALSE, xlim=c(0, top), space=0, horiz=TRUE)

```

