0.632（その２）

2016.06.15　　　株式会社　応用数理研究所　　佐々木俊久

問題

正整数nが与えられたときに，以下の式が成立する最小の正整数pを求める。

$$\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{n-p+1}\geq 1$$

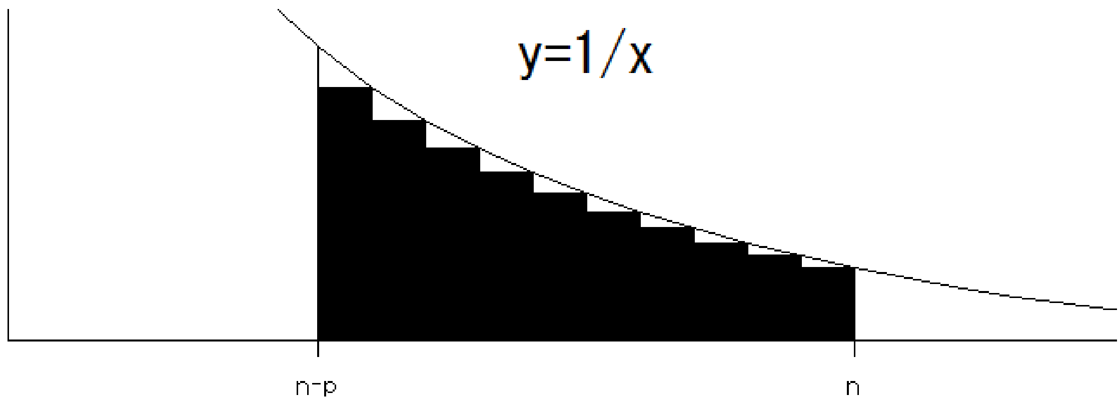
特にnを大きくした場合の近似値解

関数$y=\dfrac{1}{x}$の，区間$[n-p,n]$の積分を考えるとつぎの関係式が成立します。

$$\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{n-p+1}\leq\int_{n-p}^{n}\frac{1}{x}dx$$

右辺は，連続した区間の積分値です。

左辺は，xが$n-p+1$からnまで 1 ずつ増加したときに，幅が 1 で高さが$1/x$の長方形の面積の和です。



右辺の積分は，

$$\int_{n-p}^{n}\frac{1}{x}dx=[\log\ x]_{n-p}^{n}=\log\ n-\log\ (n-p)=\log\left(\frac{n}{n-p}\right)$$

これより，pの近似値は，

$$\log\left(\frac{n}{n-p}\right)=1$$

を解いて，

$$p = n\left(1 - \frac{1}{e}\right)$$

となります。

$$1 - \frac{1}{e} = 1 - 0.368 = 0.632$$

参考文献

事例＋演習で学ぶ　機械学習　　速水　悟著　　森北出版株式会社

9.4.3 バランスアルゴリズムの競合比

ここに，理論的にも，バランスアルゴリズムの競合比は$1-\frac{1}{e}$であることが知られています。　と記述されています。

R(S-PLUS)でのシミュレーション

次の関数(fun)により， p と p/n を求める。

```
fun <- function(n)
{
        x <- 1./n + 1./(n - 1)
        for(p in 3:n) {
                x <- x + 1./(n - p + 1)
                if(x >= 1.)
                        break
        }
        list(p = p, hi= p/n)
}

> fun(100)
$p:
[1] 64
$hi:
[1] 0.64
```

```
> fun(1000)
$p:
[1] 633
$hi:
[1] 0.633

> fun(10000)
$p:
[1] 6322
$hi:
[1] 0.6322
```

n が大きくなるほど，0.632 に近づきます。

以上