Iskanje biserov: Verjetnostne funkcije

A. Blejec

26. marec 2013

Tabla pred nekaj tedni ...

$$P(\text{za } 80 \text{ listov} < 350 \text{ potopov}) =$$

$$= P(\underbrace{v\ 349 \text{ potopih}}_{\cdot 1} \text{ vzj. } 80 \text{ listov})$$

$$P(Y \geq 80)$$

$$= P(80 \leq Y < \infty) =$$

$$= \Phi(\infty) - \Phi\left(\frac{80 - 69{,}8}{7{,}47}\right) = \frac{1}{2} - \Phi(\underline{1{,}37}) = 0{,}5 -$$

($69{,}8$ je $7{,}1 \cdot 9{,}8$)

Tabla pred nekaj tedni ...

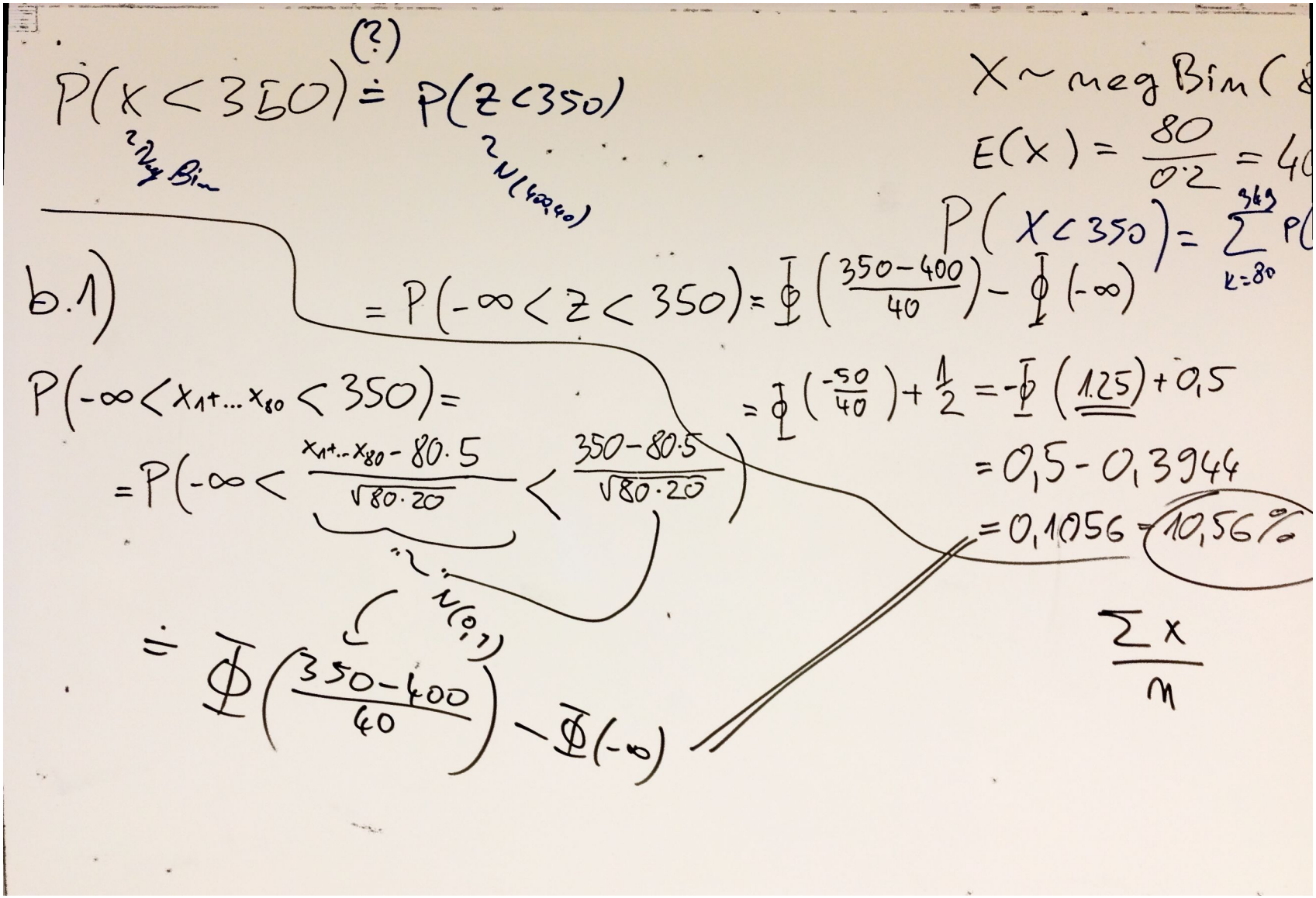
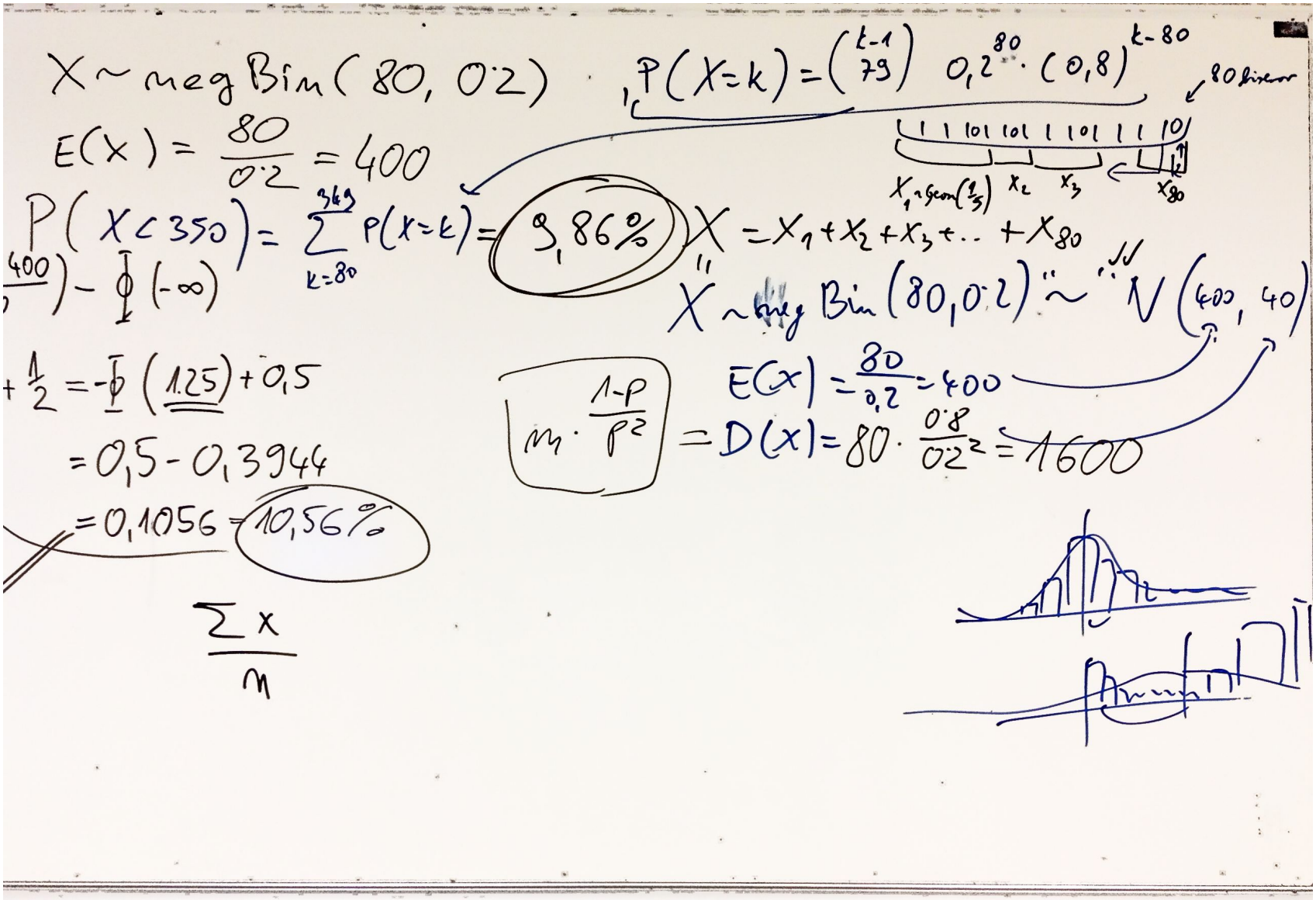


Tabla pred nekaj tedni ...



Negativna binomska porazdelitev

$$X \sim negBin(n = 80, p = 0.2)$$

$$P(X = k) = \binom{k-1}{n-1} p^n \cdot (1-p)^{k-n}$$

```
> n <- 80
> k <- 350
> p <- 0.2
> choose(k - 1, n - 1) * p^n * (1 - p)^(k - n)
[1] 0.004890492
```

V R je funkcija parametrizirana na število neuspehov.

```
> dnbinom(k - n, n, p)
[1] 0.004890492
> pnbinom(k - n, n, p) - pnbinom(k - n - 1, n, p)
[1] 0.004890492
```

Normalna aproksimacija

$$X \sim negBin(n = 80, p = 0.2)$$

$$E(X) = \frac{n}{p} = \frac{80}{0.2} = 400$$

$$D(X) = \frac{80 * 0.8}{0.2^2} = 1600$$

$$X \sim N(400, 40^2)$$

$$P(X \leq 350)$$

```
> pnorm(350, 400, 40)
[1] 0.1056498
```

Geometrijska porazdeitev

```
> n <- 80
```

$$X_k \sim geom(p = 0.2),\ k = 1, 2, \ldots, 80$$

Koliko poskusov do prvega uspeha

$$P(X = k) = (1 - p)^{k-1} * p, k = 1, 2, \ldots$$

```
> set.seed(789)
> X <- rgeom(n, p = 0.2) + 1
> X[1:5]
```

```
[1]  1 20 16  5  5
```

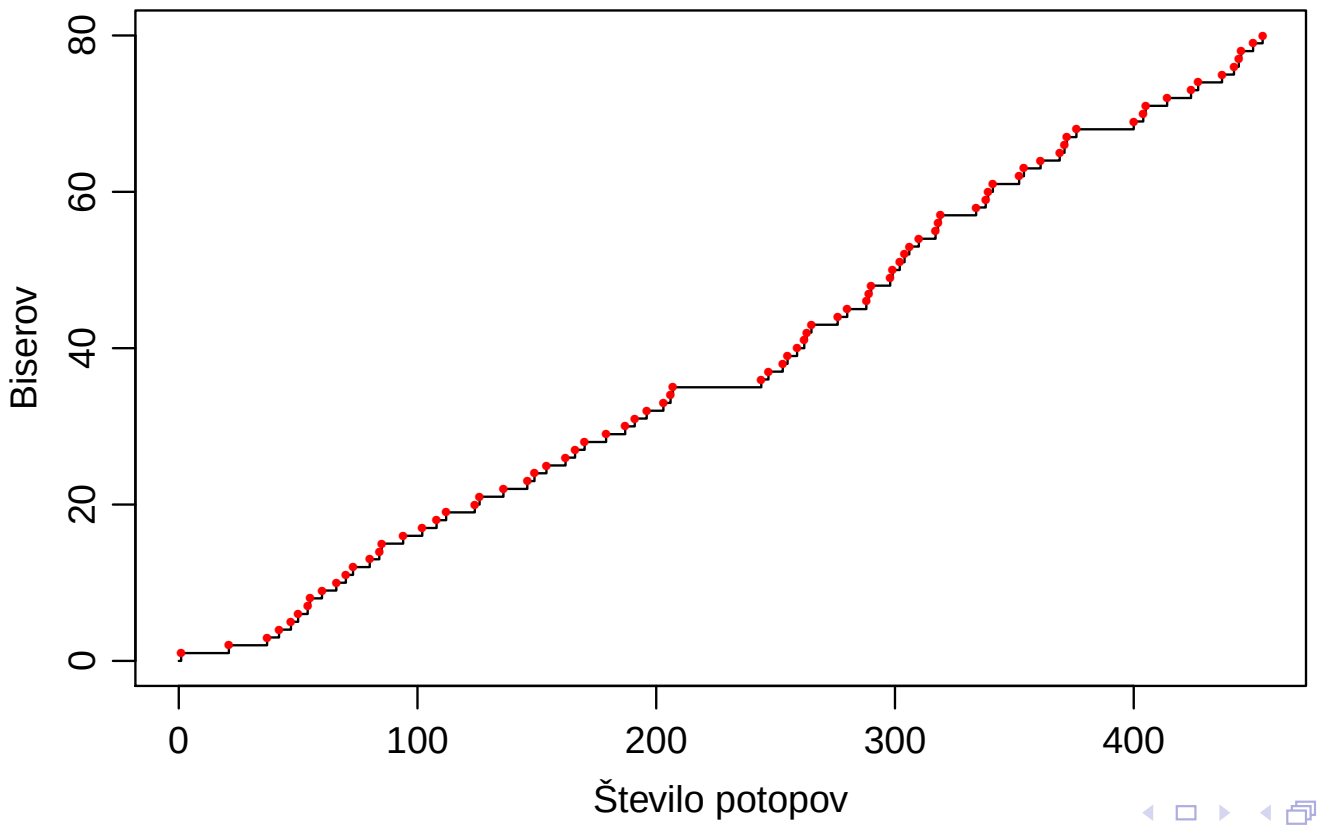
```
> (Y <- sum(X))
```

```
[1] 454
```

Zakaj $+1$? Poglejte v pomoč funkcije `rgeom`.

Potapljaški poskus

```
> par(mar = c(3, 3, 2, 1), mgp = c(2, 0.7, 0))
> X <- c(0, X)
> plot(cumsum(X), 0:n, type = "s", ylab = "Biserov",
+      xlab = "Število potopov")
> points(cumsum(X[-1]), 1:n, col = "red", pch = 16,
+      cex = 0.5)
```



Simulacija

```
> n <- 350
> skoljka <- c("prazna", "biser")
> izid <- sample(skoljka, n, replace = TRUE)
> table(izid)

izid
 biser prazna
   177    173
```

Simulacija z neenako verjetnostjo

```
> n <- 350
> skoljka <- c("prazna", "biser")
> izid <- sample(skoljka, n, replace = TRUE, prob = c(8,
+     2))
> (t <- table(izid))

izid
 biser prazna
    83    267

> t["biser"]/sum(t)

    biser
0.2371429
```

Drug pristop

```
> p0 <- 0.2
> izid <- runif(n) <= p0
> (t <- table(izid))

izid
FALSE  TRUE
  296    54

> t[2]/sum(t)

     TRUE
0.1542857
```

Z logičnimi vrednostmi lahko računamo:

```
> sum(izid)

[1] 54

> sum(izid)/length(izid)

[1] 0.1542857
```

Veliko število potapljačev N

```
> N <- 100
> n <- 350
> p <- 0.2
> X <- matrix(runif(N * n) <= p, N, n)
> X[1:4, 1:10] + 0
```

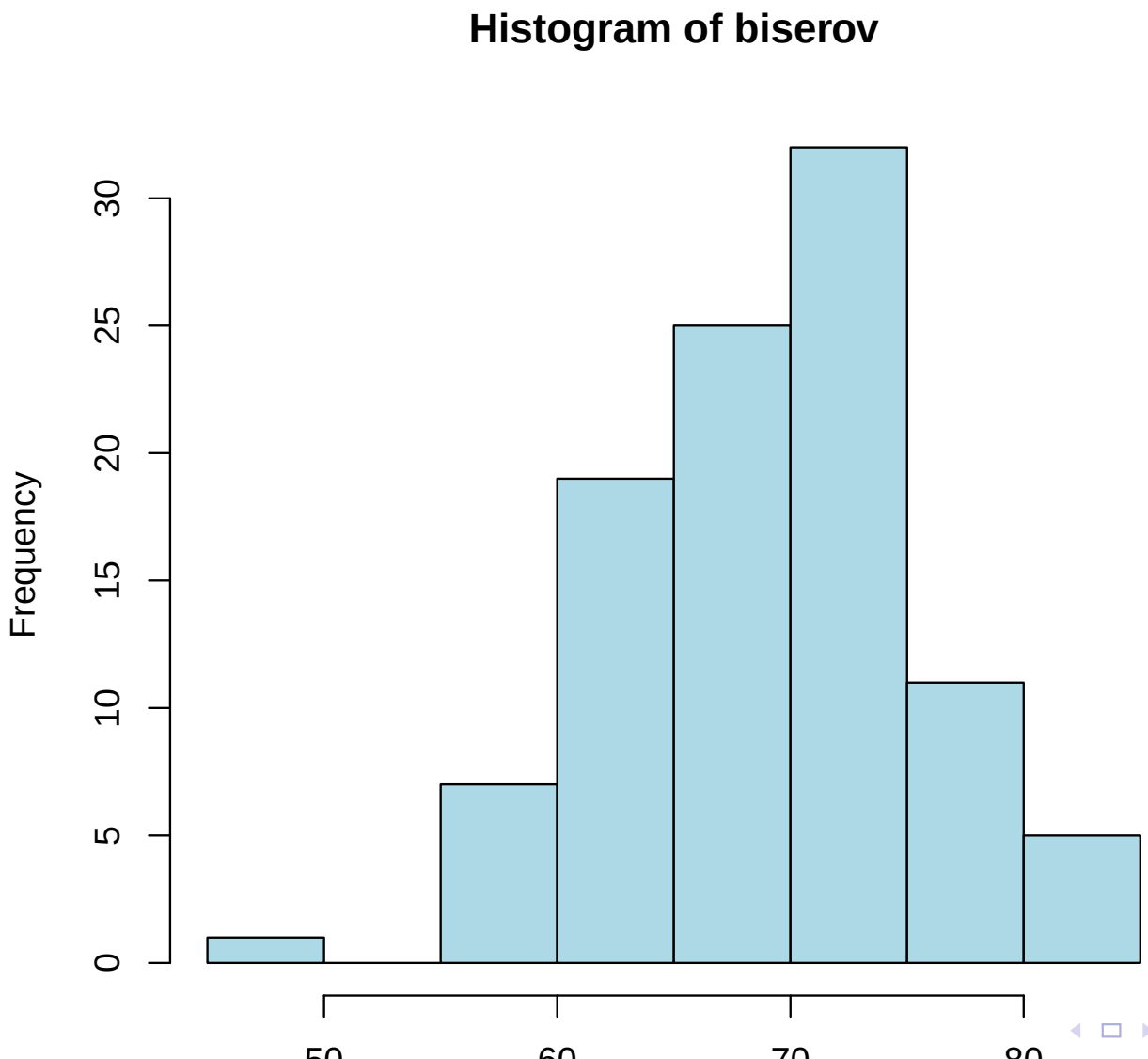
```
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
[1,]    1    0    0    0    0    0    0    0    0     0
[2,]    0    1    0    0    0    0    0    1    1     0
[3,]    1    1    0    1    0    0    0    0    0     0
[4,]    0    0    1    0    0    0    1    0    0     0
```

```
> biserov <- apply(X, 1, sum)
> head(biserov)
```

```
[1] 66 57 56 76 65 66
```

Porazdelitev

```
> hist(biserov, col = "lightblue")
```



Število potopov za 80 biserov

Z vsoto geometrijskih ali pa z negativno binomsko porazdelitvijo

```
> N <- 10000
> n <- 80
> p <- 0.2
> X <- matrix(rgeom(N * n, p = p), N, n) + 1
> X[1:4, 1:10]
```

```
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
[1,]    9    3    6    8    3    2   10    4    4     3
[2,]    3    3   10    2    7    9    1    4    1     1
[3,]    3    7    8    1    1    4    1    4    7    18
[4,]    2    1    2    2    3    7    2    7    4     1
```

```
> potopov <- apply(X, 1, sum)
> head(potopov)
```

```
[1] 422 367 408 369 393 350
```

Porazdelitev števila potopov

```
> p <- sum(potopov <= 350)/length(potopov)
> hist(potopov, col = "lightblue", main = p)
> abline(v = 350, col = "red")
```