## Moodle Tasks (Discrete Random Variables)

**Задача 1**

Хвърлят се 30 зара.

```
> dice <- sample(x = 1:6, size = 30, replace = TRUE); dice
 [1] 4 6 5 2 5 2 6 2 6 5 4 6 6 4 5 6 5 3 2 5 3 2 2 3 4 2 2 2 3 6
```

Оценете вероятността да се паднат по-малко от 5 шестици?

$$Bi\left(30, \frac{1}{6}\right)$$

```
> x <- rbinom(1000, 30, 1/6)
> sum(x <= 4) / length(x)
[1] 0.43
```

Каква е вероятността да се паднат по-малко от 5 шестици?

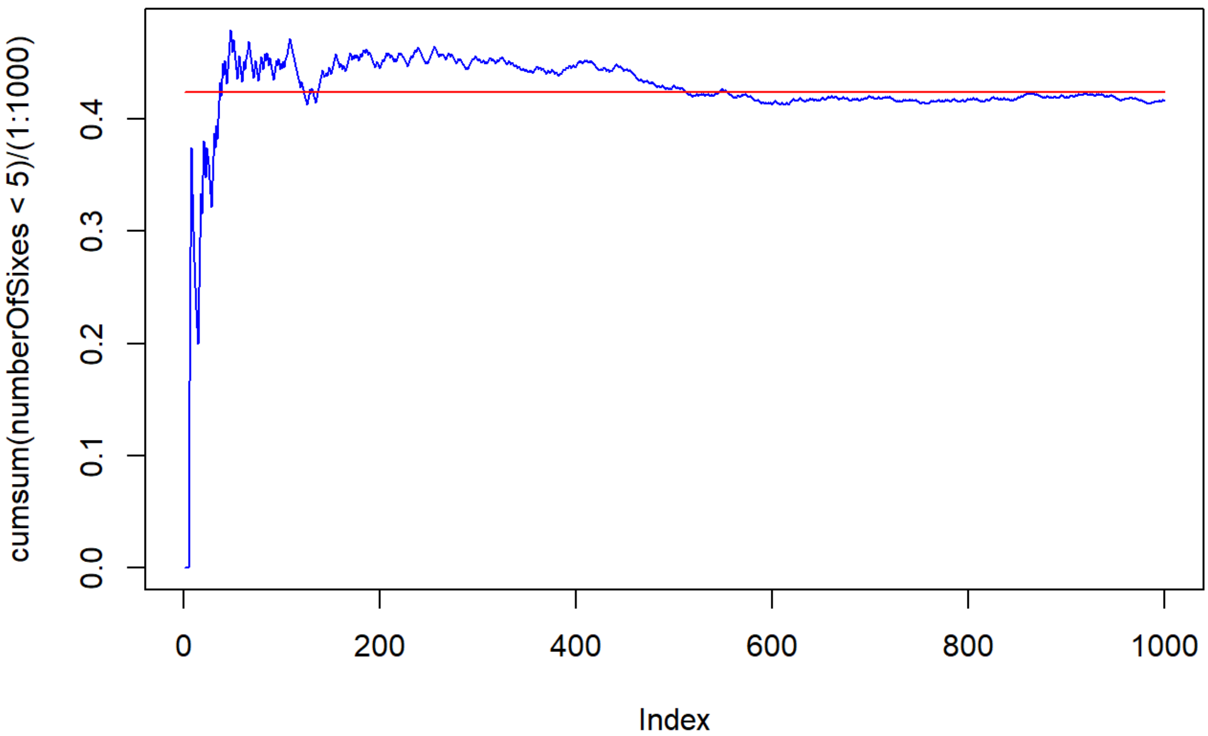
$$\mathbb{P}(X < 5) = \mathbb{P}(X \leq 4)$$

```
> pbinom(q = 4 , size = 30, prob = 1/6)
[1] 0.4243389
```

Сравнете теоретичната вероятност с експериментални данни. Виждаме, че при 1000 опита числата са много близки.

От закона за големите числа (ЗГЧ) - при увеличаване на броя на опитите оценката се приближава все повече до оценявания параметър. В случая относителната честота на опитите, при които имаме по-малко от 5 шестици все повече се приближава до вероятността на събитието да имаме по-малко от 5 шестици.

```
> numberOfSixes <- rbinom(n = 1000, size = 30 , prob = 1/6)
> plot(cumsum(numberOfSixes < 5) / (1:1000), type = "l", col = "blue")
> expected <- pbinom(4, size = 30, prob = 1/6); expected
[1] 0.4243389
> lines(rep(expected, 1000), col = "red")
```




Можем да твърдим, че с вероятност 0.75 ще се паднат повече от максимално колко шестици?

$$\max\{x \in \mathbb{R} : \mathbb{P}(X > x^*) \geq 0.75\}$$

```
> qbinom(p = 0.75, size = 30, prob = 1/6, lower.tail = FALSE)
[1] 4
```

**Задача 2**

Стрелец уцелва мишена с вероятност 0,2. За да спечели стрелецът трябва да направи три точни попадения. Каква е вероятноста за това да са му необходими:

$$X \in NegBi(n = 3, p = 0.2)$$

a) точно 8 изстрела;

$$\mathbb{P}(X = 8 - 3) = \mathbb{P}(X = 5) = \mathbb{P}(X \leq 5) - \mathbb{P}(X \leq 4)$$

```
> dnbinom(5, size = 3, prob = 0.2)
[1] 0.05505024
> pnbinom(5, size = 3, prob = 0.2) - pnbinom(4, size = 3, prob = 0.2)
[1] 0.05505024
```

b) повече от 6 изстрела;

$$\mathbb{P}(X > 6 - 3) = \mathbb{P}(X > 3)$$

```
> pnbinom(3 , size = 3, prob = 0.2, lower.tail = FALSE)
[1] 0.90112
```

c) между 5 и 8 изстрела, включително?

$$\mathbb{P}(5 - 3 \leq X \leq 8 - 3) = \mathbb{P}(2 \leq X \leq 5) = \mathbb{P}(X \leq 5) - \mathbb{P}(X \leq 1)$$

```
> pnbinom(5 , size = 3, prob = 0.2) - pnbinom(1 , size = 3, prob = 0.2)
[1] 0.1758822
```

**Задача 3**

В урна има 7 бели и 6 черни топки. От урната последователно без връщане се теглят 8 топки. Нека X е броя на изтеглените бели. Направете 1000 симулации.

Без връщане => имаме зависими опити

$X \in HG(m = 7, n = 6, k = 8)$, където
$m$ – броя на белите топки в урната
$n$ – броя на черните топки в урната
$k$ – броя на топките изтеглени от урната

```
> m <- 7; n <- 6; k <- 8
> x <- rhyper(1000, m, n, k)
```

По симулациите пресметнете границите, в който се намира $X$.

```
> range(x)
[1] 2 7
```

Оценете $\mathbb{E}[X]$ и $\mathbb{D}[X]$.

```
> mean(x)
[1] 4.339
> var(x)
[1] 0.89097
```

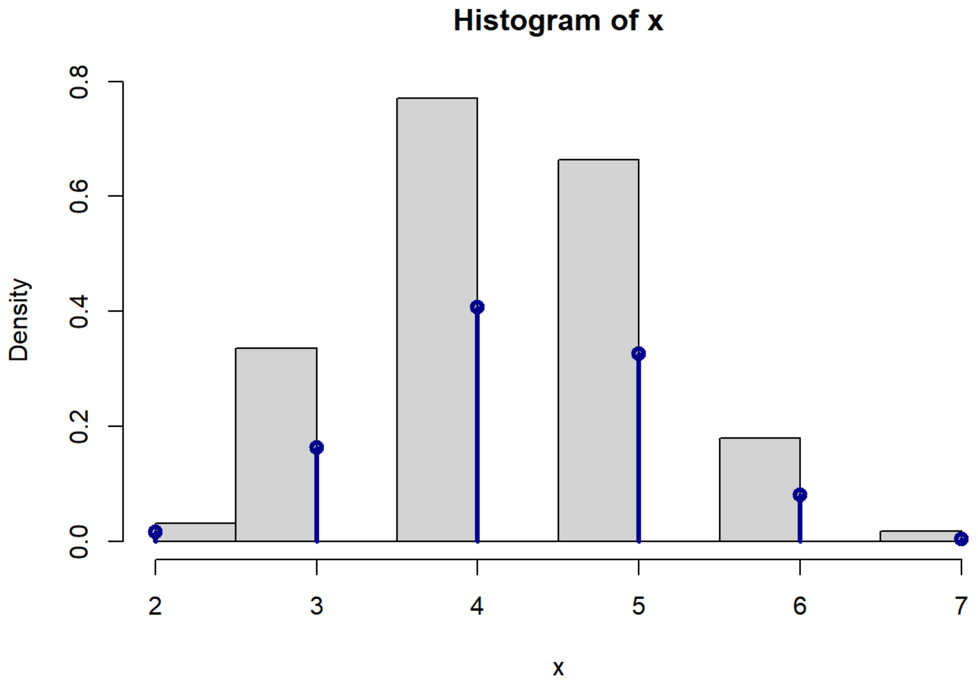
Намерете теоретичните стойности за $\mathbb{E}[X]$ и $D[X]$.

$$\mu = \mathbb{E}[X] = k\frac{m}{m+n} = 8\frac{7}{7+6} = \frac{56}{13} \approx 4.3077$$

$$\sigma^2 = \mathbb{D}[X] = \frac{kmn(m+n-k)}{(m+n)^2(m+n-1)} = \frac{8 \times 7 \times 6 \times (7+6-8)}{(7+6)^2(7+6-1)} = \frac{140}{169} \approx 0.8284$$

Представете графично емперичното и теоритичното разпределение на X (на една графика).

```
> hist(x, probability = TRUE)
> points(0:7, dhyper(0:7, m, n, k), type = "h", lwd = 3, col = "darkblue")
> points(0:7, dhyper(0:7, m, n, k), type = "p", lwd = 3, col = "darkblue")
```




**Задача 4**

Лотария се провежда със следните правила. Всеки участник избира едно число от 1 до 2n, не е необходимо да избират различни числа. Когато броят на участлиците снате n се теглят 5 печеливши числа. Каква е вероятността да се паднат тчоно две награди. Пресметнет при $n = 10, 100, 1000, 10000$. С каква случайна величина ще моделирате броя на печалбите?

**Решение:**

Имаме $5$ печеливши и $2n - 5$ губещи. От тях теглим $n$ с връщане.

$$X \in Bi\left(n, \frac{5}{2n}\right)$$

Каква е вероятността да се панат точно 2 награди е същото като $\mathbb{P}(X = 2)$. Тогава за $n = 10, 100, 1000, 10000$ ще имаме

```
> n <- c(10, 100, 1000, 10000)
> dbinom(2, n, 5 / (2*n))
[1] 0.2815676 0.2587841 0.2567403 0.2565381
```

**Задача 5**

Решете задача $4$, ако лицата теглеха различни числа.

**Решение:**

Имаме $5$ печеливши и $2n - 5$ губещи. От тях теглим $n$ без връщане.

$$X \in HG(5, 2n - 5, n)$$

Каква е вероятността да се панат точно $2$ награди е същото като $\mathbb{P}(X = 2)$. Тогава за $n = 10, 100, 1000, 10000$ ще имаме

```
> n <- c(10, 100, 1000, 10000)
> dhyper(2, 5, 2*n - 5, n)
[1] 0.3482972 0.3156646 0.3128129 0.3125313
```

**Задача 6**

За коледно парти всеки от n (n = 20) участници носи по един подарък. Подаръците се номерират и в шапка се слагат номерата от 1 до n. Всеки участник си тегли номер и получава съответния подарък. Напишете функции, които пресмятат:

a) теоретичната вероятност;
b) емперичната вероятност изчислена по 10000 опита;

никой да не получи свояподарък.

Пресметнете очакването на броя хора получили своя подарък.

**Решение:**

Нека $A_1$ е събитието 1-вия да си е получил подаръка, $A_2$ е събитието 2-рия да си е получил подаръка,
…
$A_n$ е събитието n-тия да си е получил подаръка.

$$
\begin{aligned}
\mathbb{P}(\overline{A}_1 \cap \overline{A}_2 \cap \ldots \cap \overline{A}_n) = 1 - \mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) &= \\
&= 1 - \sum_{l=1}^{n} (-1)^{l+1} \sum_{1 \le i_1 < \ldots < i_l \le n} \mathbb{P}(A_{i_1} \ldots A_{i_l}) = \\
&= 1 - \sum_{l=1}^{n} (-1)^{l+1} \sum_{1 \le i_1 < \ldots < i_l \le n} \frac{1^l (n-l)!}{n!} = \\
&= 1 - \sum_{l=1}^{n} (-1)^{l+1} \binom{n}{l} \frac{(n-l)!}{n!} = \\
&= 1 - \sum_{l=1}^{n} (-1)^{l+1} \frac{n!}{l!(n-l)!} \frac{(n-l)!}{n!} = \\
&= 1 - \sum_{l=1}^{n} \frac{(-1)^{l+1}}{l!} = 1 + \sum_{l=1}^{n} \frac{(-1)^{l}}{l!} = \\
&= \sum_{l=0}^{n} \frac{(-1)^{l}}{l!} = \sum_{l=0}^{20} \frac{(-1)^{l}}{l!} = 0.3679
\end{aligned}
$$

Сега да симулираме $10\ 000$ пъти пермутациите на числата $1, \ldots, 20$.

```
> library(pracma)
> presentPermut <- replicate(10000, randperm(1:20, 20))
```

Сега да преброим в колко от тях нито един индекс не съвпада с числото

```
> a <- 0
> for (i in 1:10000) {
+   a[i] <- sum(presentPermut[, i] == 1:20) == 0
+ }
> sum(a)
[1] 3654
```

Да сметнем емперичната вероятност и да я сравним с теоретичната получена по-горе.

```
> sum(a) / 10000
[1] 0.3654
```