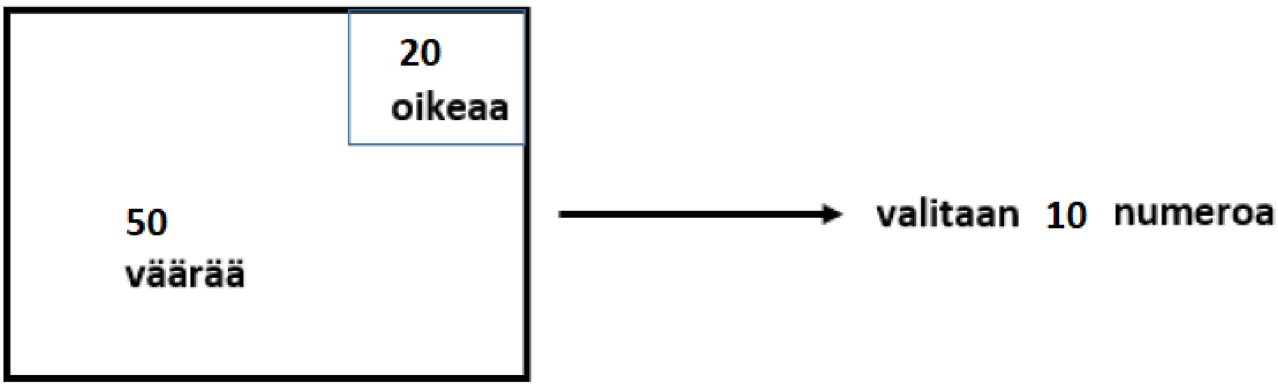
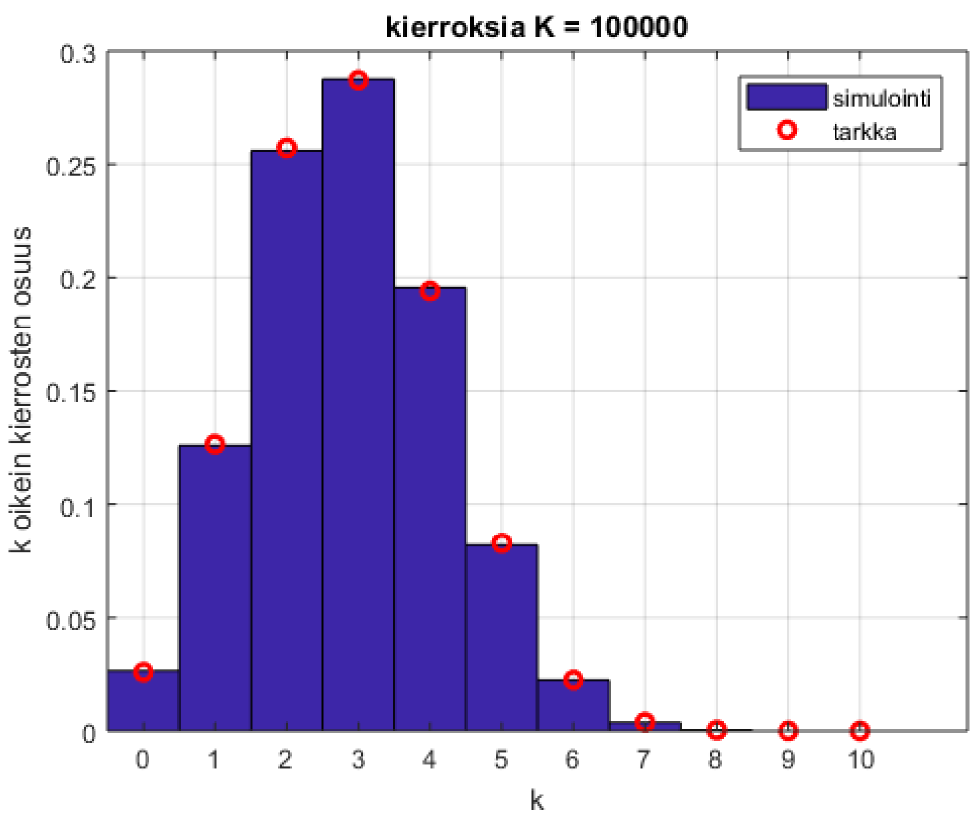
1. Keno: player chooses 10 numbers from 1,2,...,70. There are 20 winning and 50 non-winning numbers. Calculate the probabilities that player chooses 0,1,...,10 winning numbers and test by simulation.



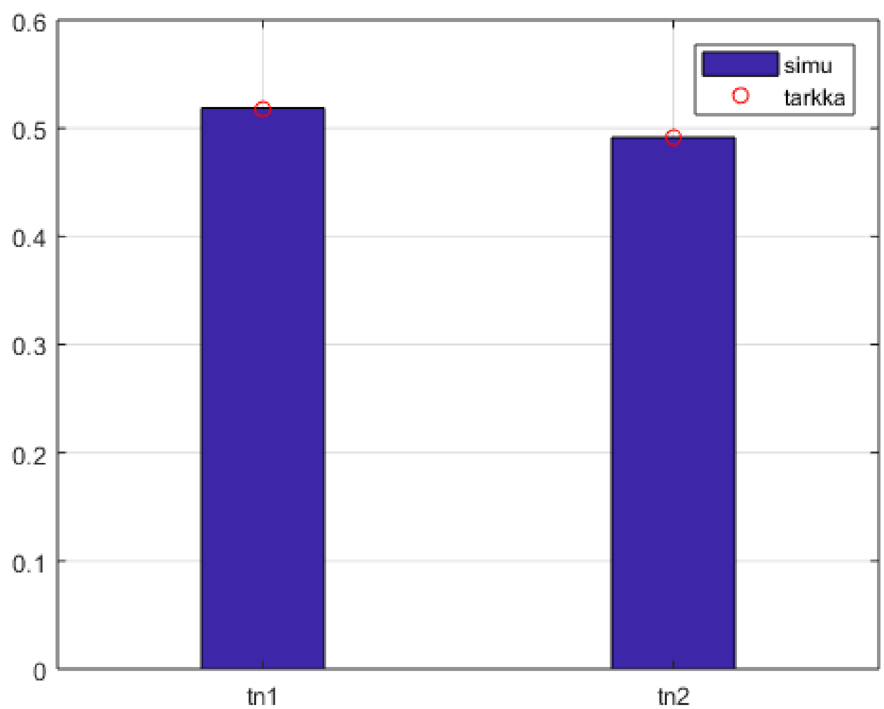


2. (De Mere problems) Calculate probabilities that

1) by throwing a dice 4 times at least on 6 occurs

2) by throwing two dice 24 times at least one double 6 occurs

Test by simulation
hint: result1=randi(6,1,4), result2=randi(6,2,24)
sum(M) is a row vector containing the sums of the columns of a matrix M



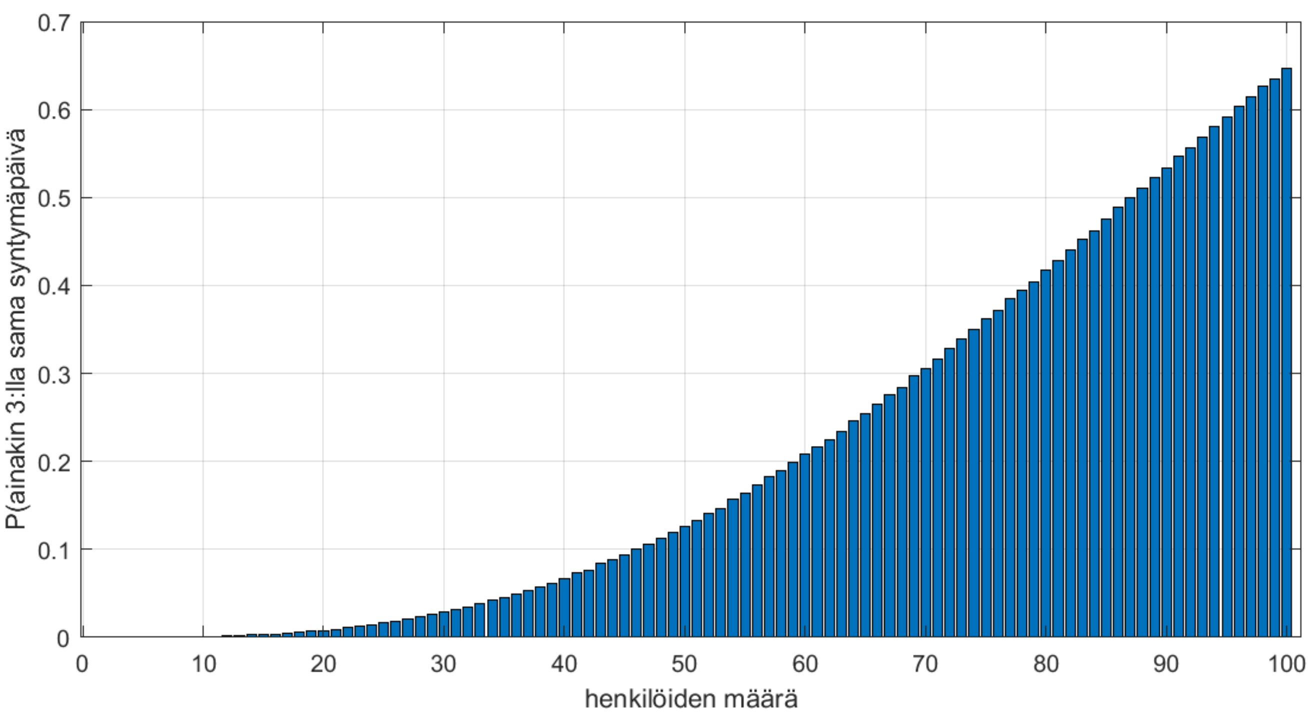
3. In lotto 7 numbers from 1,2,...,39 are chosen. Calculate by simulation the probability there are no consecutive numbers

ans: probability $= \binom{33}{7}/\binom{39}{7} \approx 0.28$

hint: sort the numbers using command sort

4. Calculate by simulation the probility that n birthdays contain all days of the year. How large n should be in order that the probability is > 0.5 ? (ans. 2287)

5. Calculate by similation the probability that among n birthdays at least 3 are same



hint: command mode might be useful

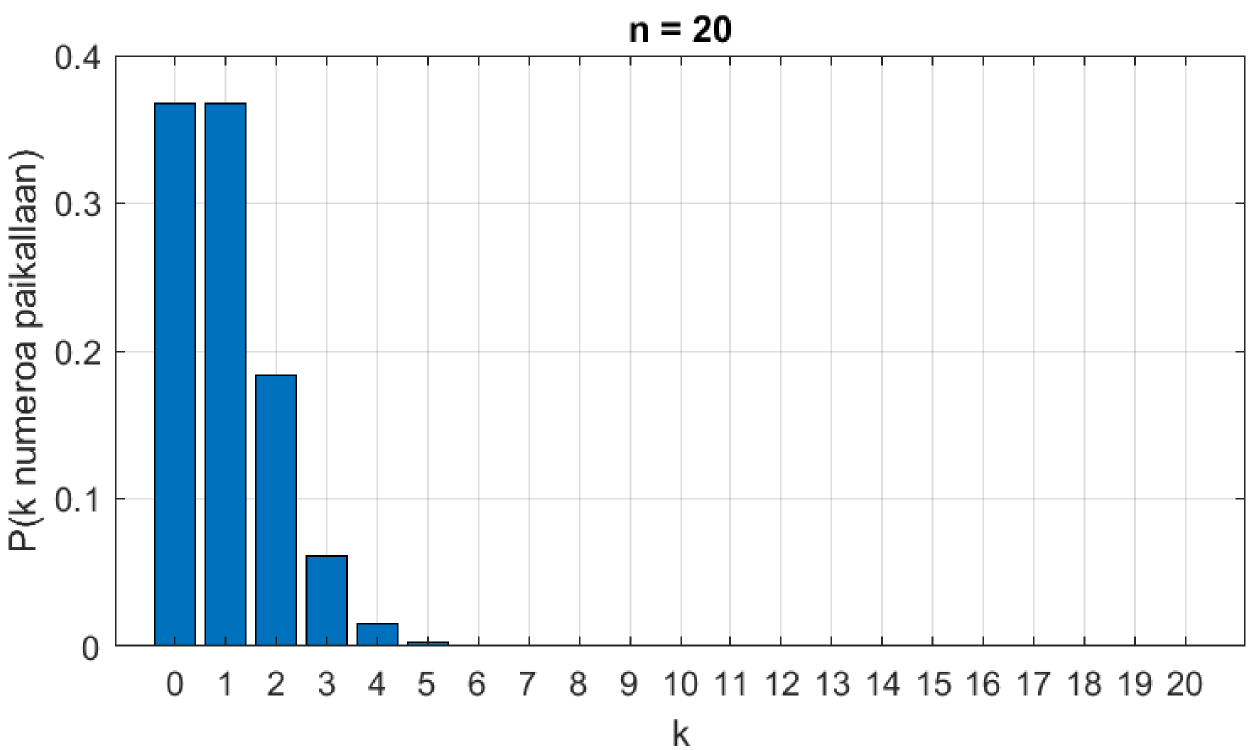
6. Deal 52 cards to 4 persons, 13 to each. Calculate the probabilities

$$P(\text{all aces (numbers 1) go to same person})$$

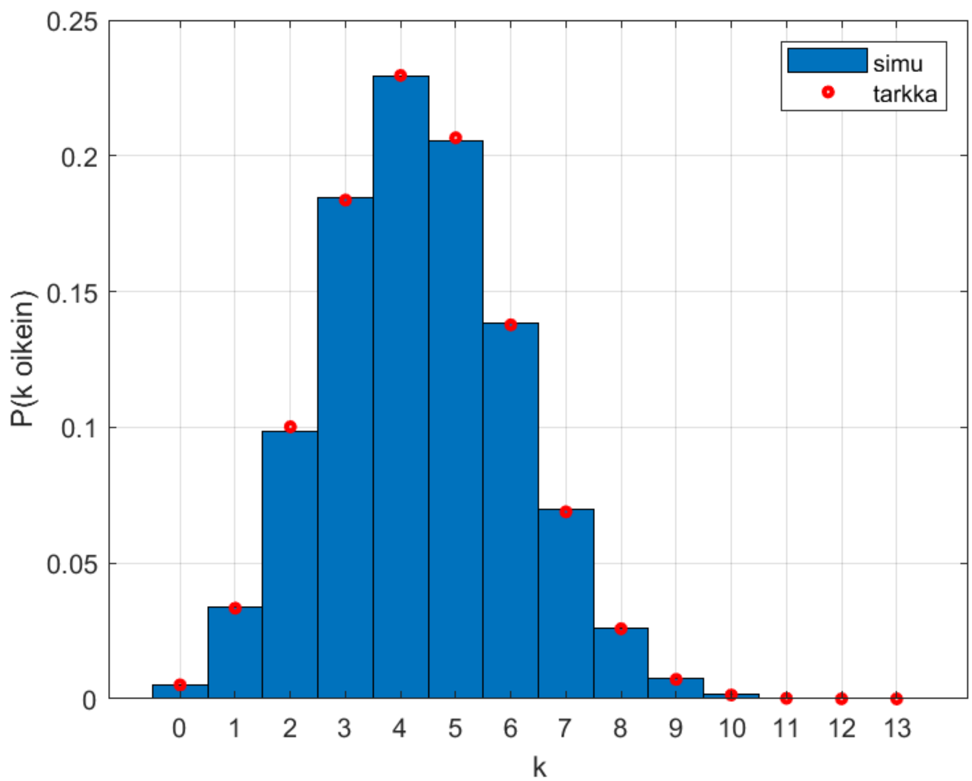
$$P(\text{each person gets one ace})$$

Test by simulation. (vast: 0.0106/0.1055)

7. Arrange numbers $1, 2, \ldots, n$ to a random aorder. Calculate the probability that exactly $k = 0, 1, \ldots, n$ numbers stay on their original places. Test by simulation.

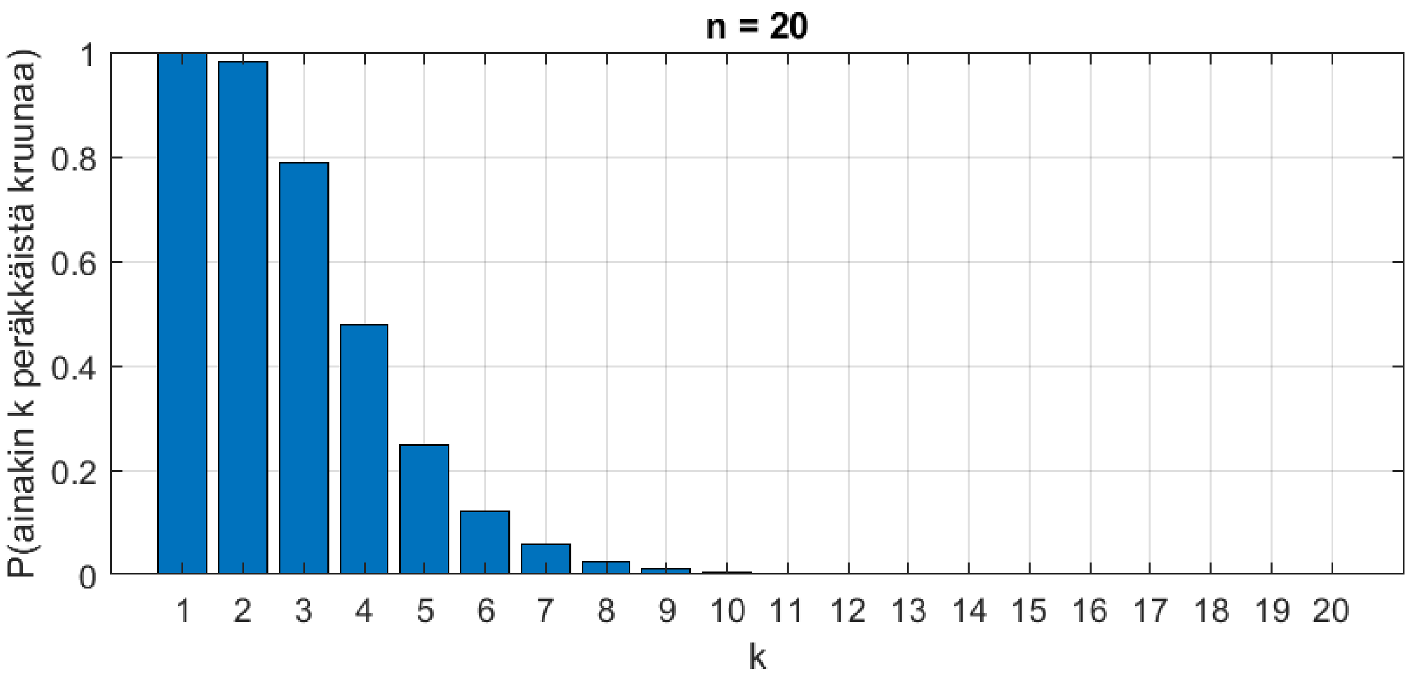


8. In vakio player choose one of 3 possibility for the result of each of 13 games. Calculate the probabilities that he gets $k = 0, 1, 2, \ldots, 13$ games right. Test by simulation



hint: if vector u contains the results and v players guesses, command sum$(u == v)$ might be useful

9. Calculate by simulation the probability that when flipping a coin n times at least $k = 1, 2, ..., n$ consecutive heads appear



10. (Coupon collector's problem) In probability theory, the coupon collector's problem describes "collect all coupons and win" contests. It asks the following question: If each box of a brand of cereals contains a coupon, and there are n different types of coupons, what is the probability that more than t boxes need to be bought to collect all n coupons?
Calculate by simulation the probability that by bying t boxes one finds all n different coupons.

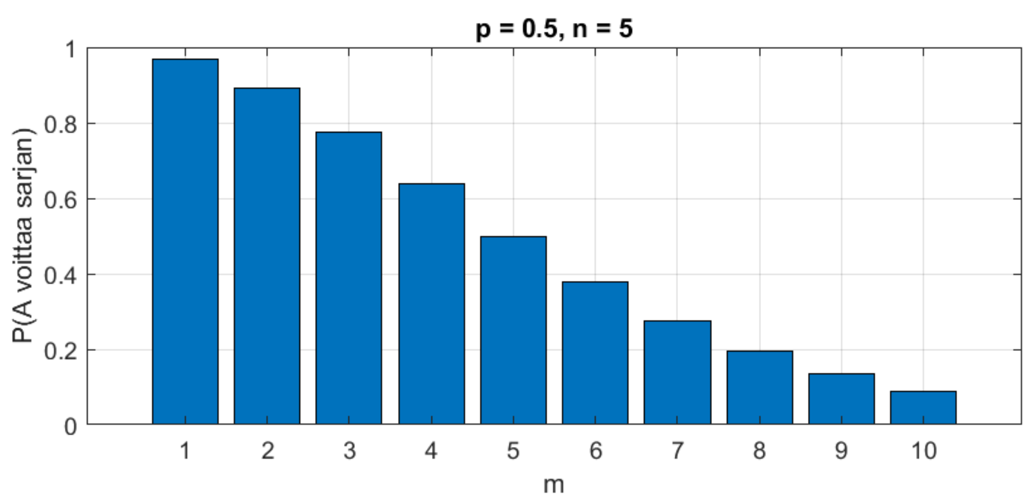
vast: $n = 10, t = 20/30/40/50 \rightarrow 0.215/0.63/0.86/0.95$

11. Problem of points

A and B play a game which A wins with probability p and B with probability $1-p$. They continue until one of them has won N games. Calculate by simulation that A wins N games first, if A has $N-m$ and B has $N-n$ wins (i.e A wins m games before B wins n games).

hint: probability is

$$P_{m,n} = \sum_{r=m}^{m+n-1} \binom{m+n-1}{r} p^r (1-p)^{m+n-1-r}$$



12. Box I contains b_1 black and w_1 white balls, box II b_2 black and w_2 white and box III b_3 black and w_3 white. Choose a box and from there a ball randomly. Calculate the probability that the ball is black. Test by simulation.

ans:

$b_1 = 5, w_1 = 7, b_2 = 3, w_2 = 4, b_3 = 4, w_3 = 6 \rightarrow P(\text{black}) = 0.415$

$b_1 = 5, w_1 = 7, b_2 = 3, w_2 = 4, b_3 = 4, w_3 = 6 \rightarrow P(\text{black}) = 0.608$

13. Craps. Player twrows two dice. If the sum is 7 or 11, player wins, and if the sum is 2,3 or 12, player loses. If sum is 4,5,6,8,9 or 10, player continues throwing the two dice until sum is the same as in the first throw and player wins, or the sum is 7 and player loses. Calculate the probability that player wins. Test by simulation.

vast: 0.493

Hint:

$W =$ player wins

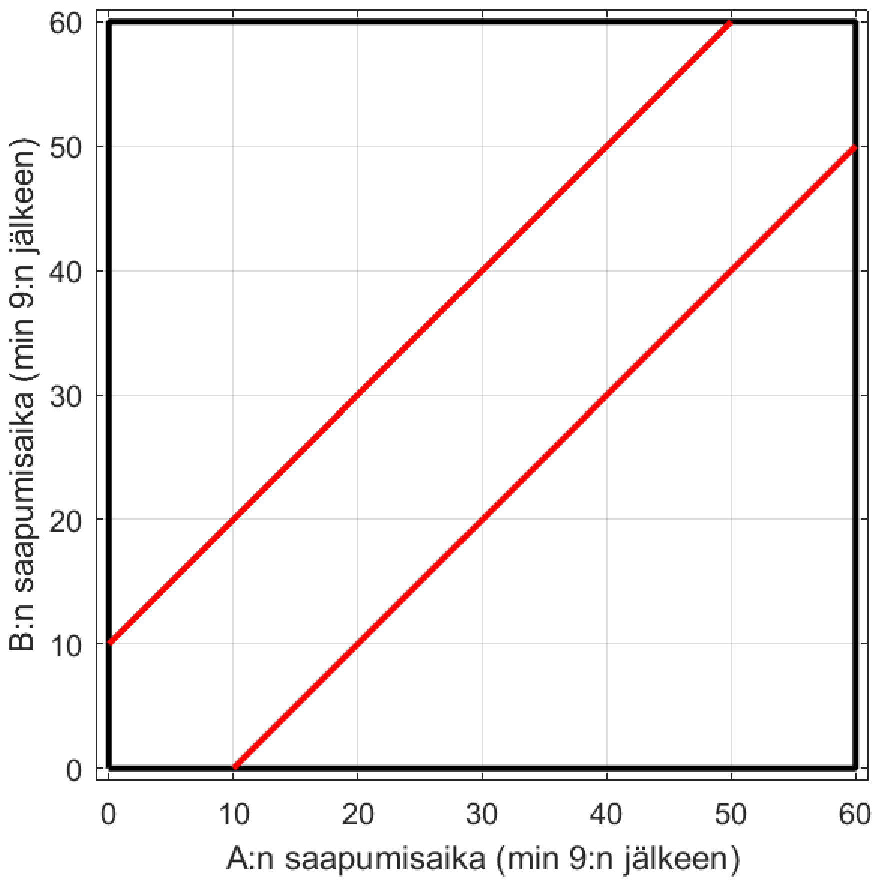
$S_k =$ first sum is k, $k = 2, 3, \ldots, 12$

$P(W) = P(S_1)P(W|S_1)+P(S_2)P(W|S_2)+\cdots+P(S_{12})P(W|S_{12})$

If $k = 4, 5, 6, 8, 9, 10$, then

$$P(W|S_k) = P(\text{sum } k \text{ before sum } 7)$$

14. A and B arrive to a cafe at random times between 9-10, and stay 10 minutes. Calculate the probability that are in the cafe simultaneously.
Test by simulation (arrival times (minutes after 9) are uniformly distributed numbers from 0...60).

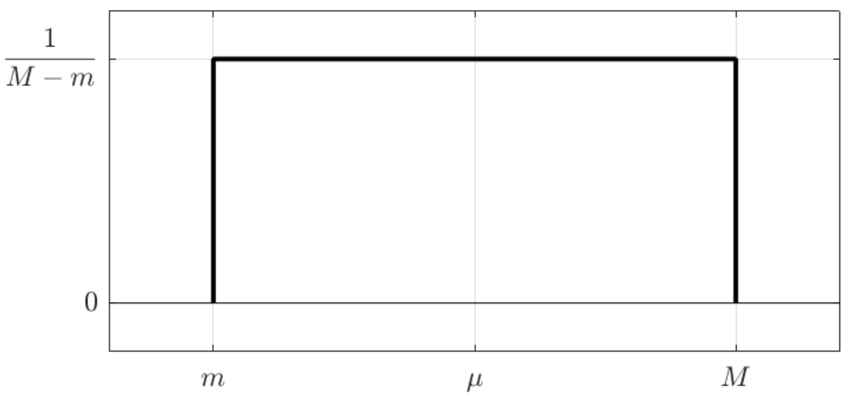


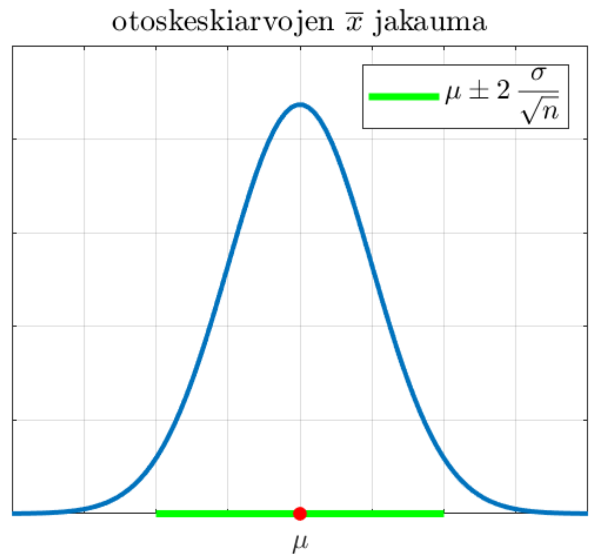
(ans: 0.305)

15. Mean $\overline{x}$ of (sufficiently many) uniformly distributed numbers $x_1, x_2, \ldots, x_n$ is approximately normally distributed, $\overline{x} \sim N(\mu, \sigma^2/n)$, where

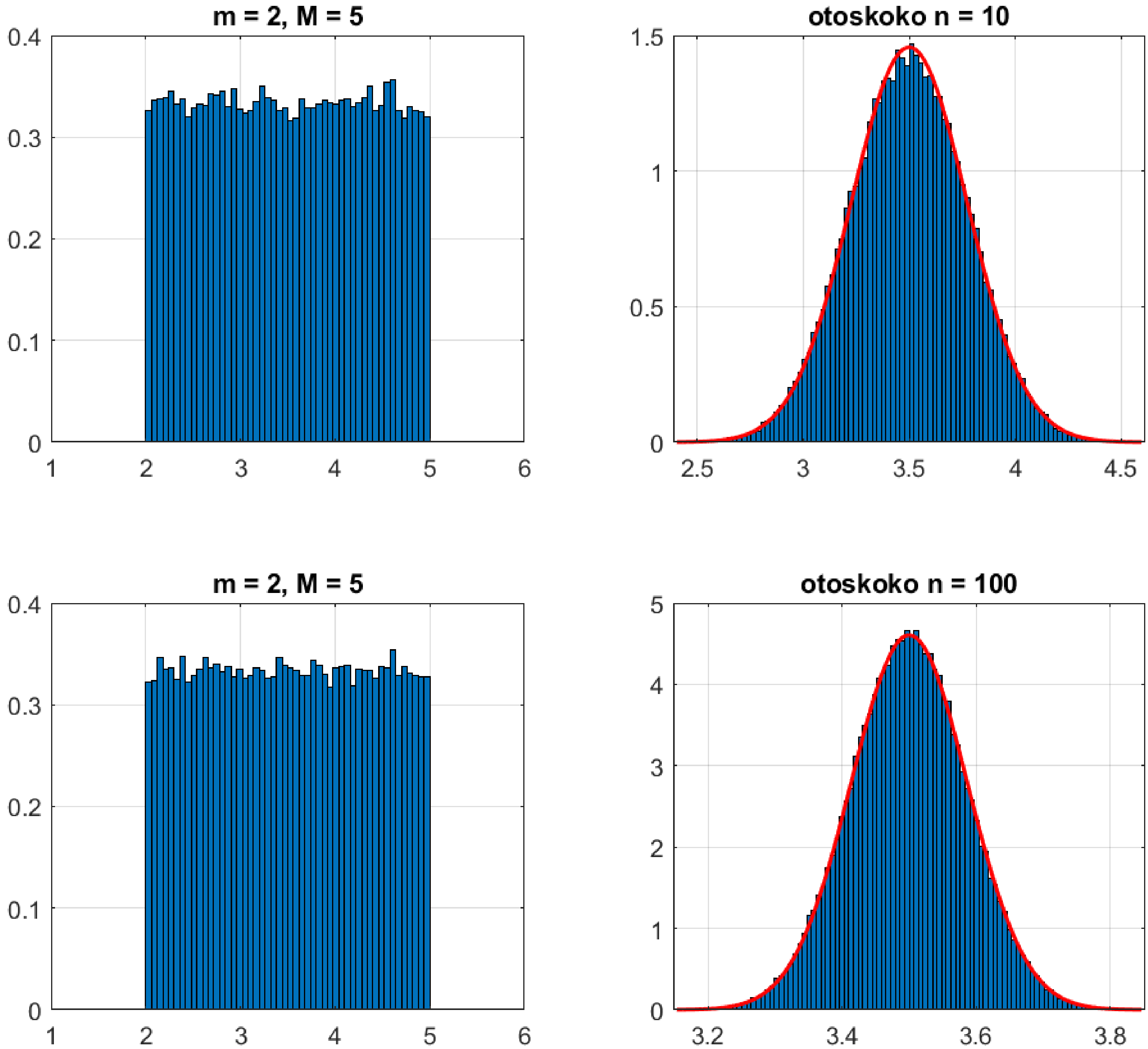
$$\mu = \frac{1}{2}(m+M), \quad \sigma = \frac{1}{\sqrt{12}}(M-m)$$

are the mean and standard deviation of uniform distribution on the interval $m \ldots M$





Test by simulation: create for example, 100000 uniformly distributed numbers, take 100000 samples of size n and draw the histogram of sample means and the $N(\mu, \sigma^2/n)$-gaussian to the same picture



m = 2, M = 5
otoskoko n = 10
m = 2, M = 5
otoskoko n = 100

16. If $x_1 \sim N(\mu_1, \sigma_1^2)$ and $x_2 \sim N(\mu_2, \sigma_2^2)$, then

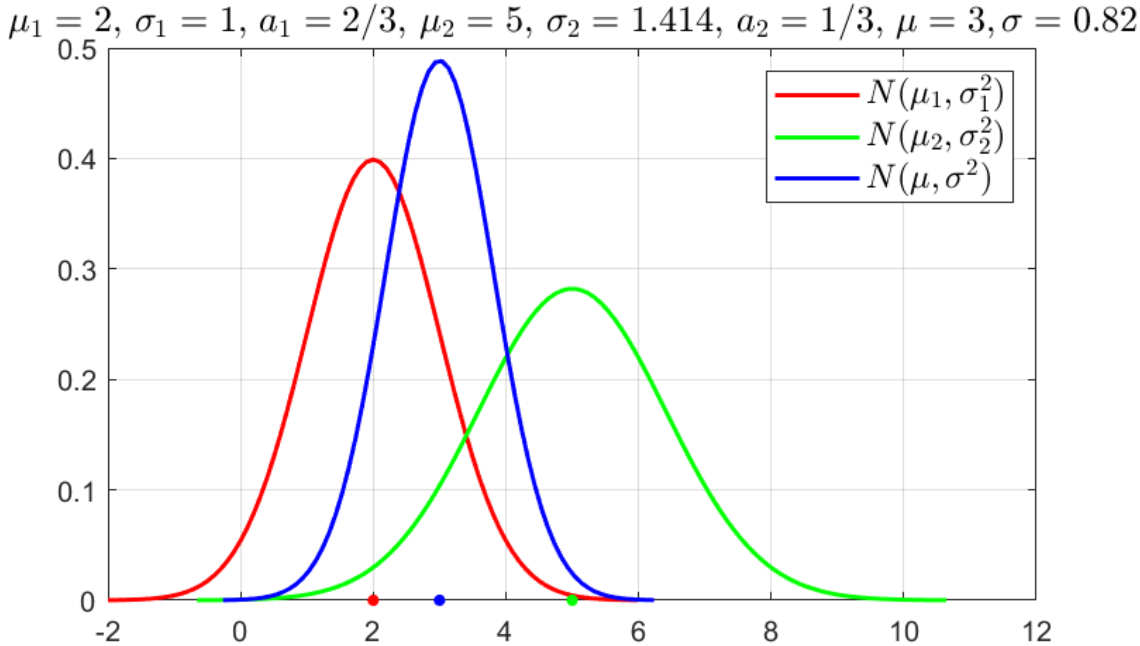
$$x = a_1 x_1 + a_2 x_2 \sim N(\mu, \sigma^2)$$

where

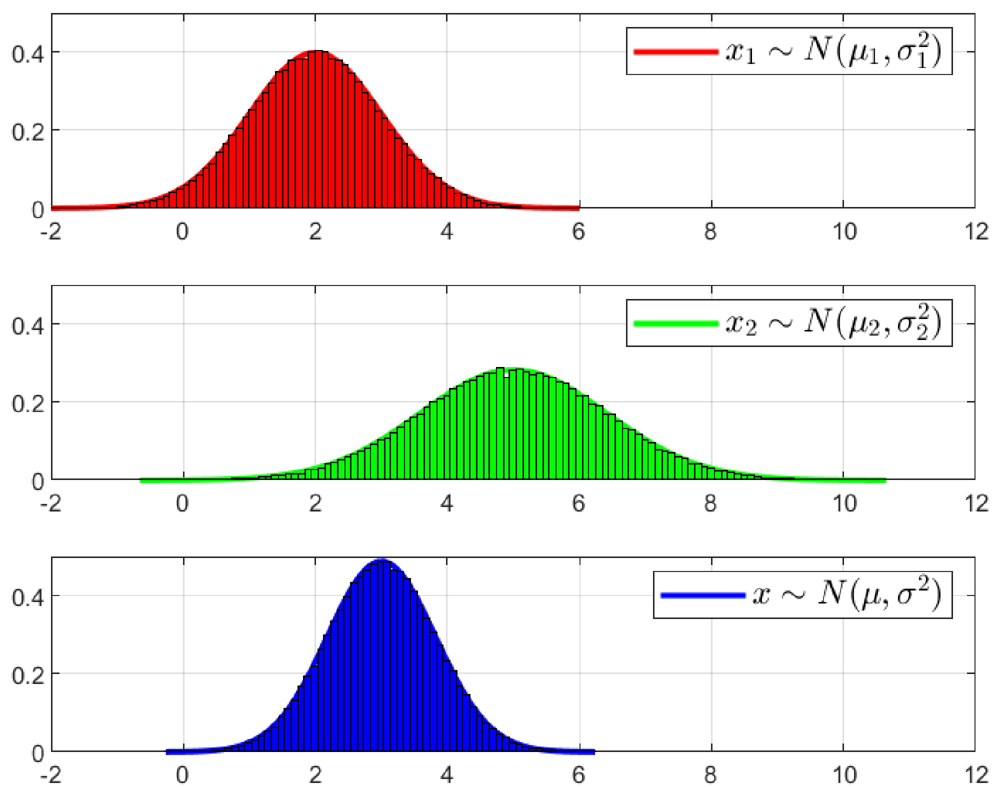
$$\mu = a_1\, \mu_1 + a_2\, \mu_2$$

$$\sigma^2 = a_1^2\, \sigma_1^2 + a_2^2\, \sigma_2^2$$

Test by simulation: given means μ_1, μ_2, standard deviations σ_1, σ_2 and coefficents a_1, a_2, create for example 100000 numbers $x_1 \sim N(\mu_1, \sigma_1^2)$ and $x_2 \sim N(\mu_2, \sigma_2^2)$, calculate $x = a_1 x_1 + a_2 x_2$ and draw the histograms and Gaussian corresponding to $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ and $N(\mu, \sigma^2)$



$\mu_1 = 2,\ \sigma_1 = 1,\ a_1 = 2/3,\ \mu_2 = 5,\ \sigma_2 = 1.414,\ a_2 = 1/3,\ \mu = 3, \sigma = 0.82$
$N(\mu_1, \sigma_1^2)$
$N(\mu_2, \sigma_2^2)$
$N(\mu, \sigma^2)$



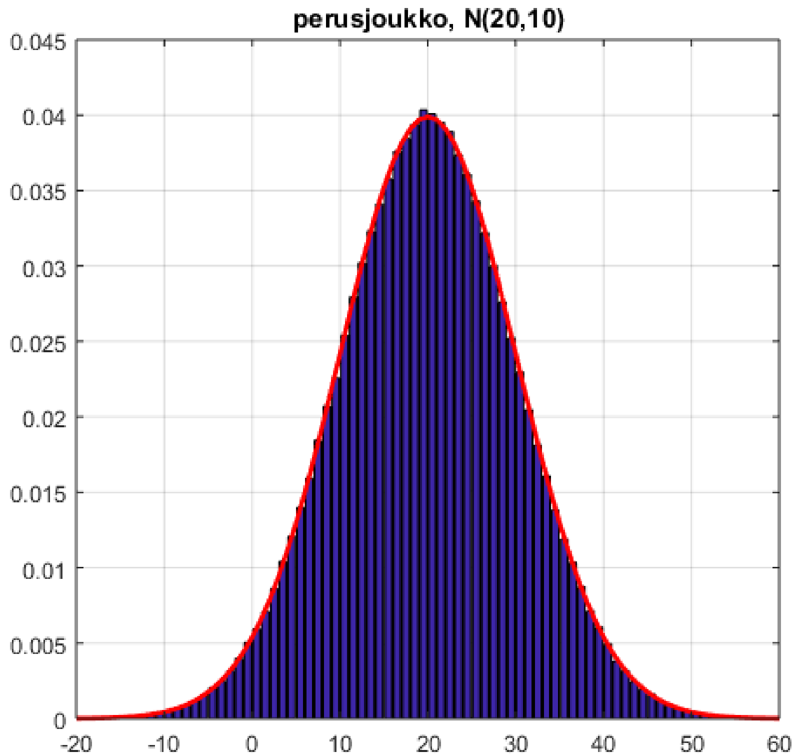
$x_1 \sim N(\mu_1, \sigma_1^2)$
$x_2 \sim N(\mu_2, \sigma_2^2)$
$x \sim N(\mu, \sigma^2)$

17. Create 100000 samples of size n from $N(\mu, \sigma^2)$, calculate sample standard deviations s and draw the histogram of

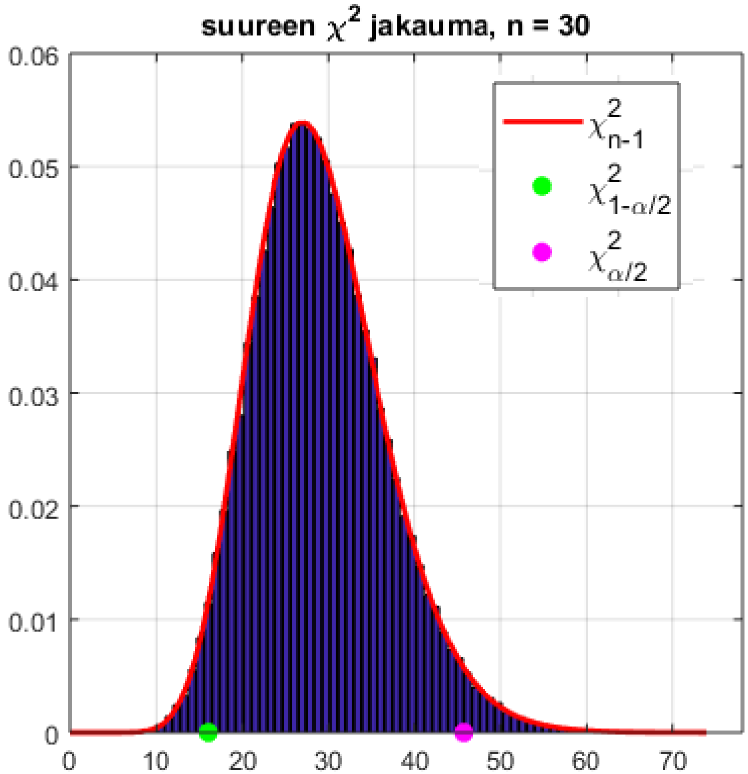
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

and the probability density function of χ^2-distribution with $n-1$ degrees of freedom. Calculate also how many of samples are such that σ is on the 95 % confidence interval

$$\sqrt{\frac{n-1}{\chi^2_{\alpha/2}}} \cdot s \quad \ldots \quad \sqrt{\frac{n-1}{\chi^2_{1-\alpha/2}}} \cdot s$$



perusjoukko, N(20,10)
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
-20
-10
0
10
20
30
40
50
60



suureen χ^2 jakauma, n = 30
χ^2_{n-1}
$\chi^2_{1-\alpha/2}$
$\chi^2_{\alpha/2}$
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
40
50
60
70

18. Values of X are **Poisson-distributed** with parameter $\lambda > 0$, if

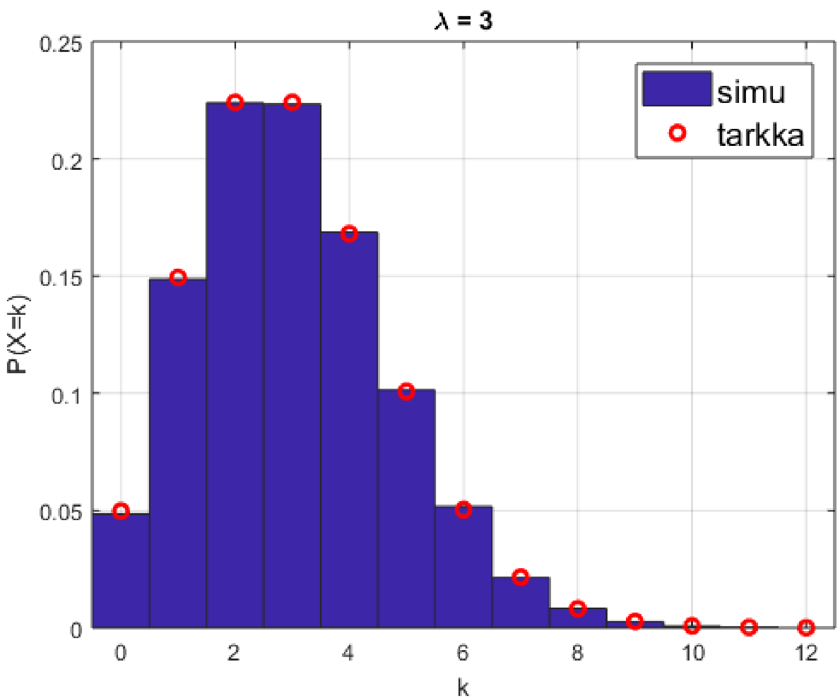
$$P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

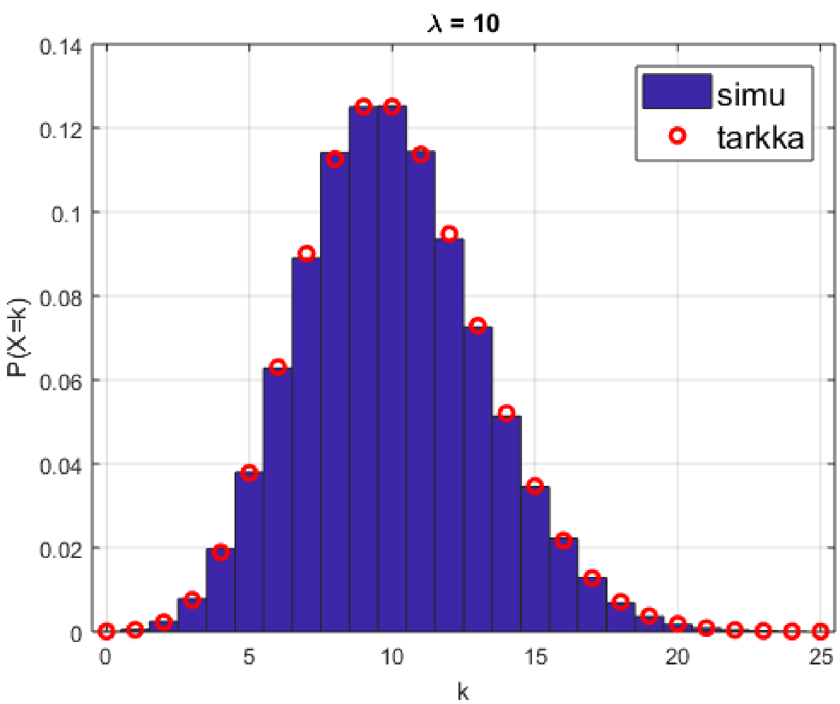
Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed unit interval of time Parameter λ is the average number of events occurring

Simulate Poisson-distribution as follows: given λ and length T of time interval, repeat for example 100000 times:

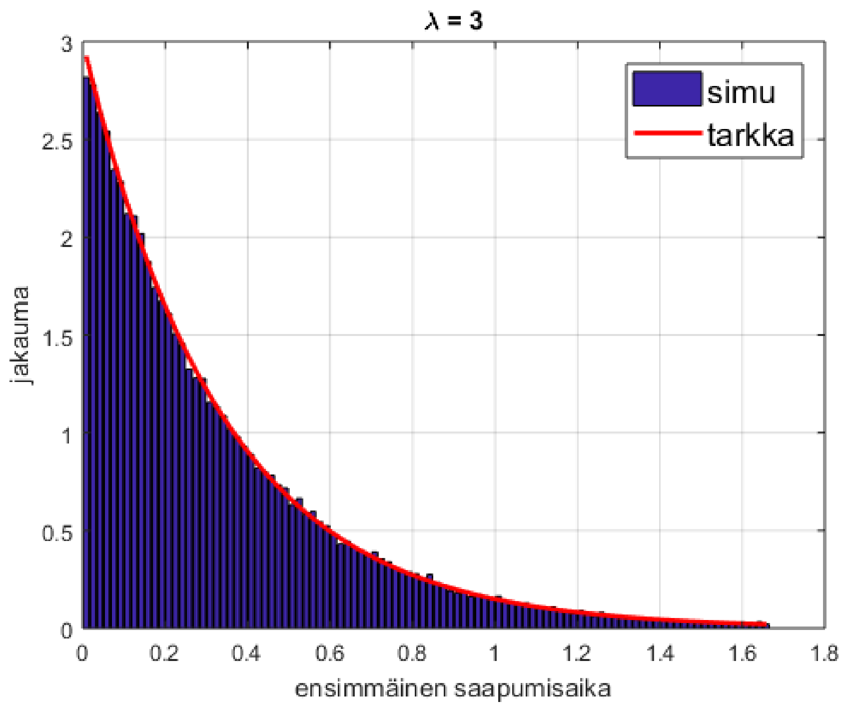
There are on the average $N = \lambda \cdot T$ ocurrences on time interval $0 \ldots T$. Create the times of occurrences from uniform distribution, and calculate how many occurences are in time interval $0 \ldots 1$

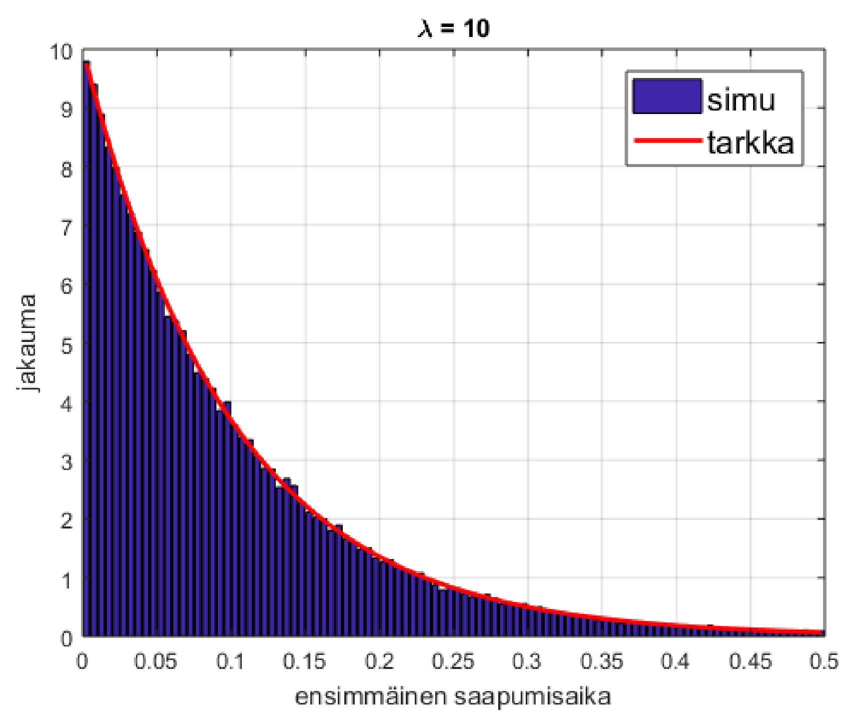
Draw a picture like below showing the probabilities $P(X = k)$, $k = 0, 1, 2, \ldots, k_{max}$, and their simulated versions
(below $T = 1000$ and $k_{max} = 12$ and 25)





19. Collect in the ex18 simulation the first occurrences and draw their histogram.



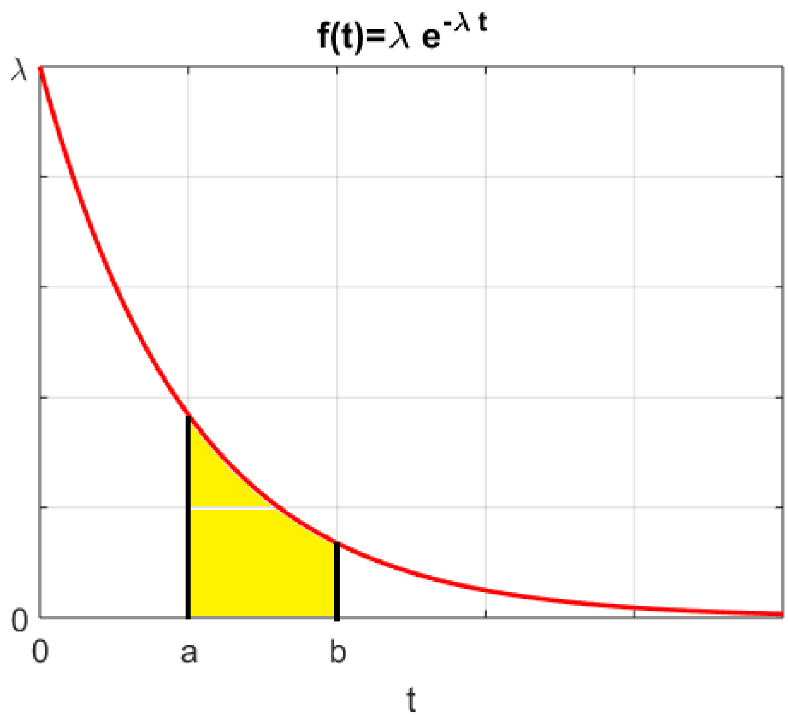


They should be approximately exponentially distributed i.e the probability that first occurrence is on the interval $a \ldots b$ is

$$\int_a^b f(t)\,dt = e^{-\lambda a} - e^{-\lambda b}$$

where the probability density function

$$f(t) = \lambda \cdot e^{-\lambda t}$$



Draw the graph of the probability density function $f(t)$ to the same picture with the histogram

20. Given the means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and covariance matrices Σ_1 and Σ_2, draw the 95 % ellipses of 2D-normal distributions $N(\boldsymbol{\mu}_1, \Sigma_1)$, $N(\boldsymbol{\mu}_2, \Sigma_2)$ and $N(\boldsymbol{\mu}, \Sigma)$ where

$$\boldsymbol{\mu} = A_1\boldsymbol{\mu}_1 + A_2\boldsymbol{\mu}_2$$

$$\Sigma = A_1\Sigma_1A_1^T + A_2\Sigma_2A_2^T$$

$$A_1 = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}\Sigma_1^{-1}$$

$$A_2 = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}\Sigma_2^{-1}$$

Test by simulation: create 100000 points

$$X_1 \sim N(\boldsymbol{\mu}_1, \Sigma_1),\ X_2 \sim N(\boldsymbol{\mu}_2, \Sigma_2)$$

calculate

$$X = A_1X_1 + A_2X_2$$

and draw a picture

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \Sigma_1 = \begin{bmatrix} 9 & 3.6 \\ 3.6 & 4 \end{bmatrix}, \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 1 & 1.8 \\ 1.8 & 9 \end{bmatrix}$$