ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Лекция 16

<u>Проверка гипотезы о виде функции распределения</u>

Для этого потребуется распределение случайной величины χ_K^2

$$\tilde{\chi}_K^2 = \sum_{k=1}^{K} x_k^{\,2}, \quad \text{где все} \ \left\langle x_k \right\rangle = 0$$

Приведение исходной случайной величины
к единичной дисперсии:

$$\xi_k^{\,2} = \frac{x_k^{\,2}}{\sigma_x^2}, \qquad \sigma_\xi^2 = 1 \qquad\qquad \chi_K^2 = \sum_{k=1}^{K} \xi_k^{\,2} = \sum_{k=1}^{N} \frac{x_k^{\,2}}{\sigma_x^2} = \frac{\tilde{\chi}_K^2}{\sigma_x^2}$$

$$w\left(\chi_K^2\right) = \frac{1}{2^{K/2}\Gamma\left(K/2\right)} \left(\chi_K^2\right)^{K/2-1} e^{-\chi_K^2/2}$$

Более общий случай

Имеется N независимых случайных величин $x_k \sim \mathbf{N}\left(\left\langle x_k \right\rangle, \sigma_{x_k}^2\right)$

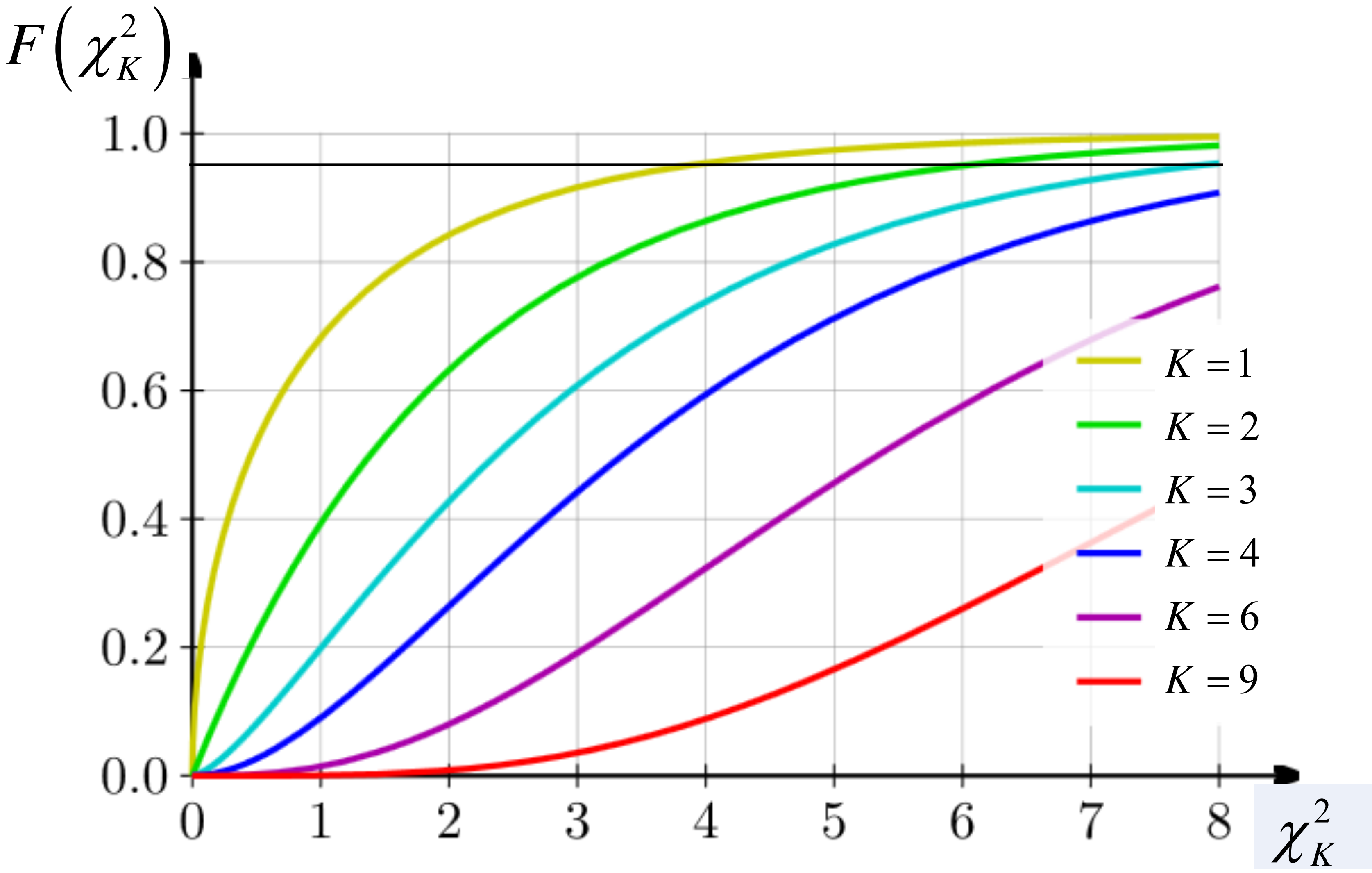
Требуется все эти величины привести
к нулевому математическому ожиданию и единичной дисперсии

$$\xi_k = \frac{x_k - \left\langle x_k \right\rangle}{\sigma_{x_k}}$$

$$\left\langle \xi_k \right\rangle = 0, \qquad \sigma_{\xi_k}^2 = 1$$

$$\chi_K^2 = \sum_{k=1}^{K} \xi_k^{\ 2}$$

Функция распределения



Пример данных

Число кавалеристов прусской армии, погибших за 20 лет из-за гибели под ними коня. Всего 200 донесений. Случайная величина – число погибших, доложенное в одном донесении.

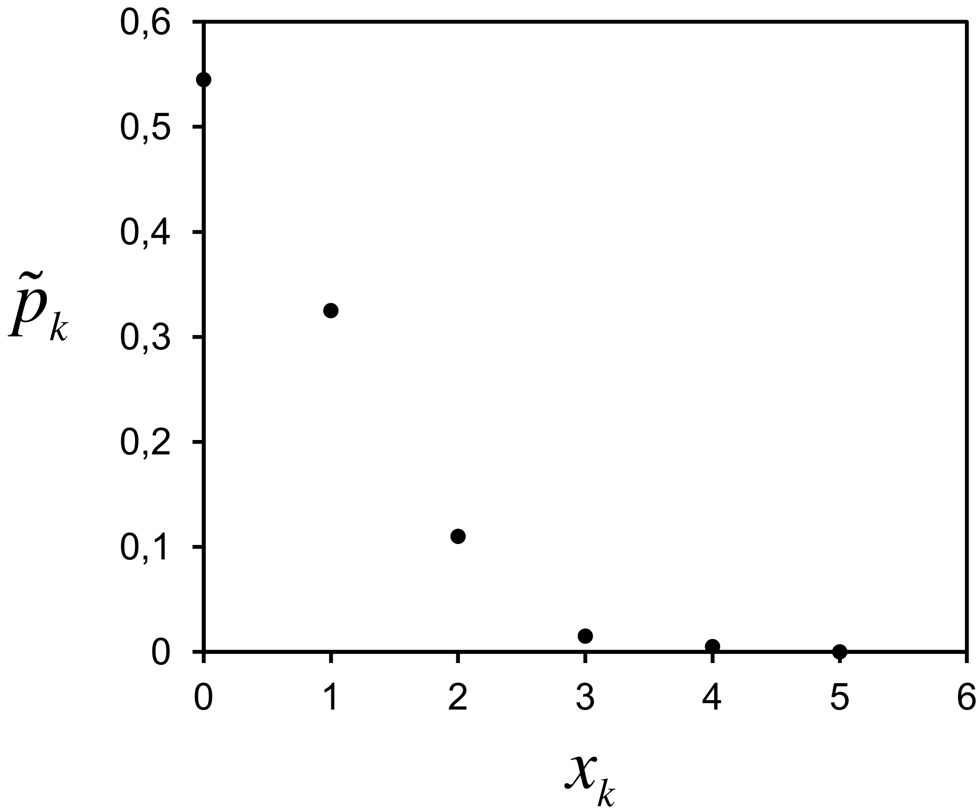
k – номер значения случайной величины, x_k – значение случайной величины, N_k – число испытаний, в результате которых получено k-е значение случайной величины, $\tilde{p}_k$ – частоты значений случайной величины

k	x_k	N_k	$n_k = x_k N_k$	$\tilde{p}_k = \dfrac{N_k}{\sum_k N_k}$
1	0	109	0	0.545
2	1	65	65	0.325
3	2	22	44	0.110
4	3	3	9	0.015
5	4	1	4	0.005
6	≥ 5	0	0	0

$$\sum_k N_k = 200 \quad \sum_k x_k N_k = 122$$

$$\sum_k N_k = 200, \quad \sum_k x_k N_k = 122, \quad \bar{x} = \frac{\sum_k x_k N_k}{\sum_k N_k} = 0.61$$

Гистограмма:



Гипотеза:

$$p_x(\beta) = \frac{\beta^x}{x!} e^{-\beta}$$

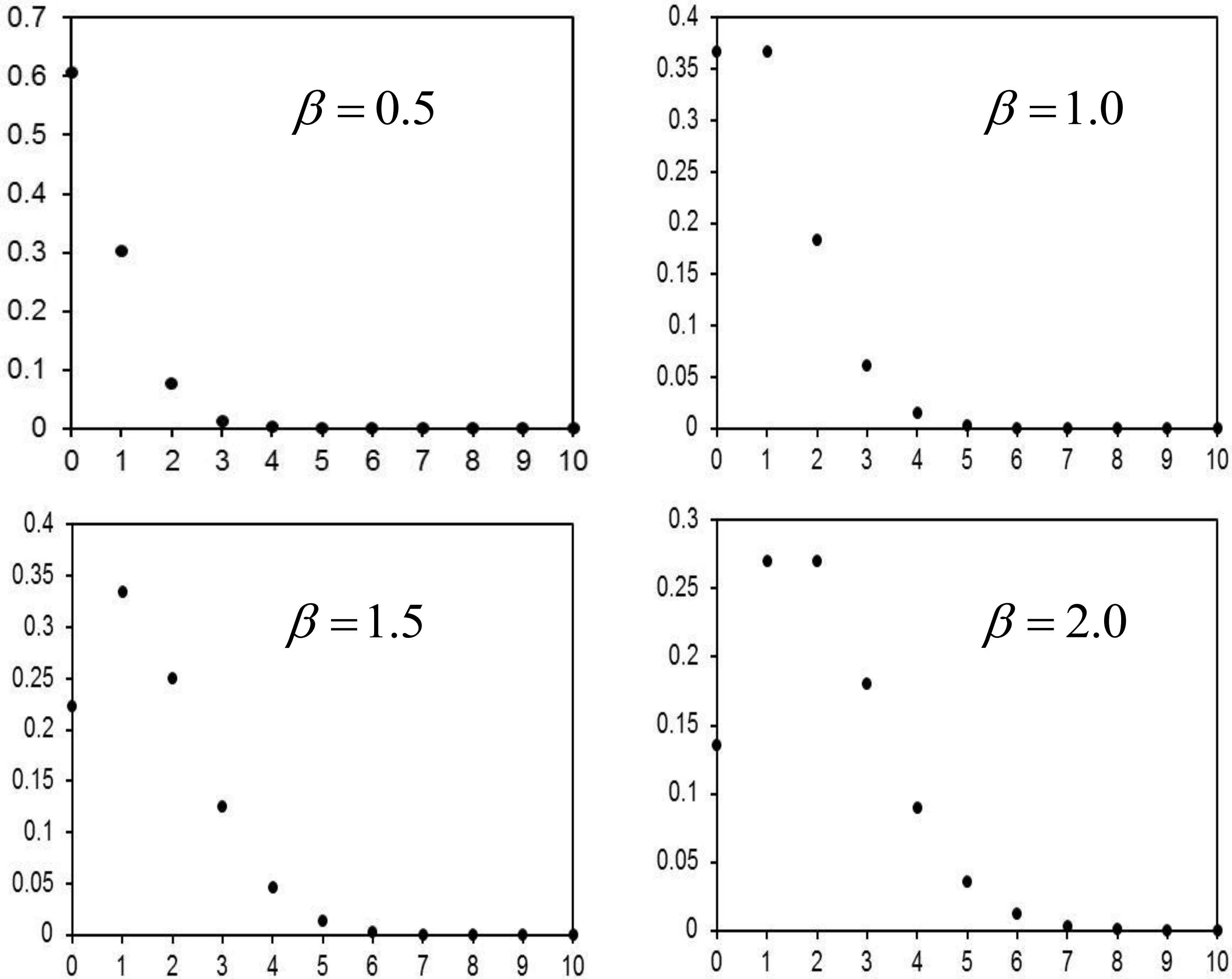
Что предположить о β ?

$$\langle x \rangle = \beta, \quad \sigma_x^2 = \beta$$

Гипотеза:

$$\beta = \bar{x} = 0.61$$

Распределение Пуассона



Каждая x_k — отдельная случайная величина

$$p_k = \frac{\beta^{x_k}}{x_k\,!} e^{-\beta}$$ — для одного испытания

Для N испытаний случайной величины x_k :

$$P_N\left(N_k\right) = \frac{N!}{N_k\,!\left(N - N_k\right)!}\left(p_k\right)^{N_k}\left(1 - p_k\right)^{N - N_k}$$

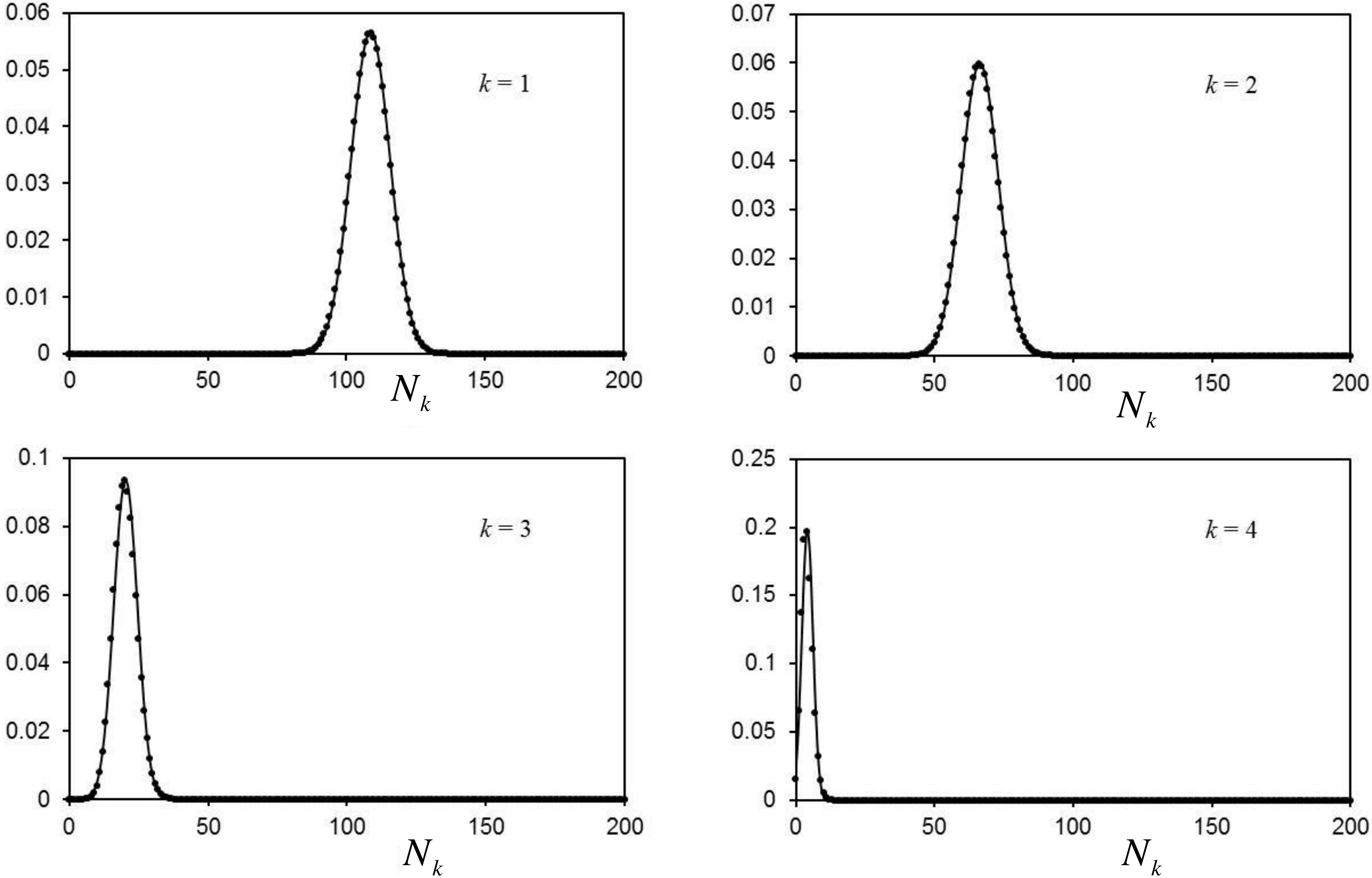
$$\left\langle N_k \right\rangle = Np_k\,, \quad \sigma_{N_k}^2 = Np_k\left(1 - p_k\right), \quad \sigma_{N_k} = \sqrt{Np_k\left(1 - p_k\right)}$$

$$\xi_k = \frac{N_k - \langle N_k \rangle}{\sigma_{N_k}} = \frac{N_k - Np_k}{\sqrt{Np_k\left(1 - p_k\right)}}$$

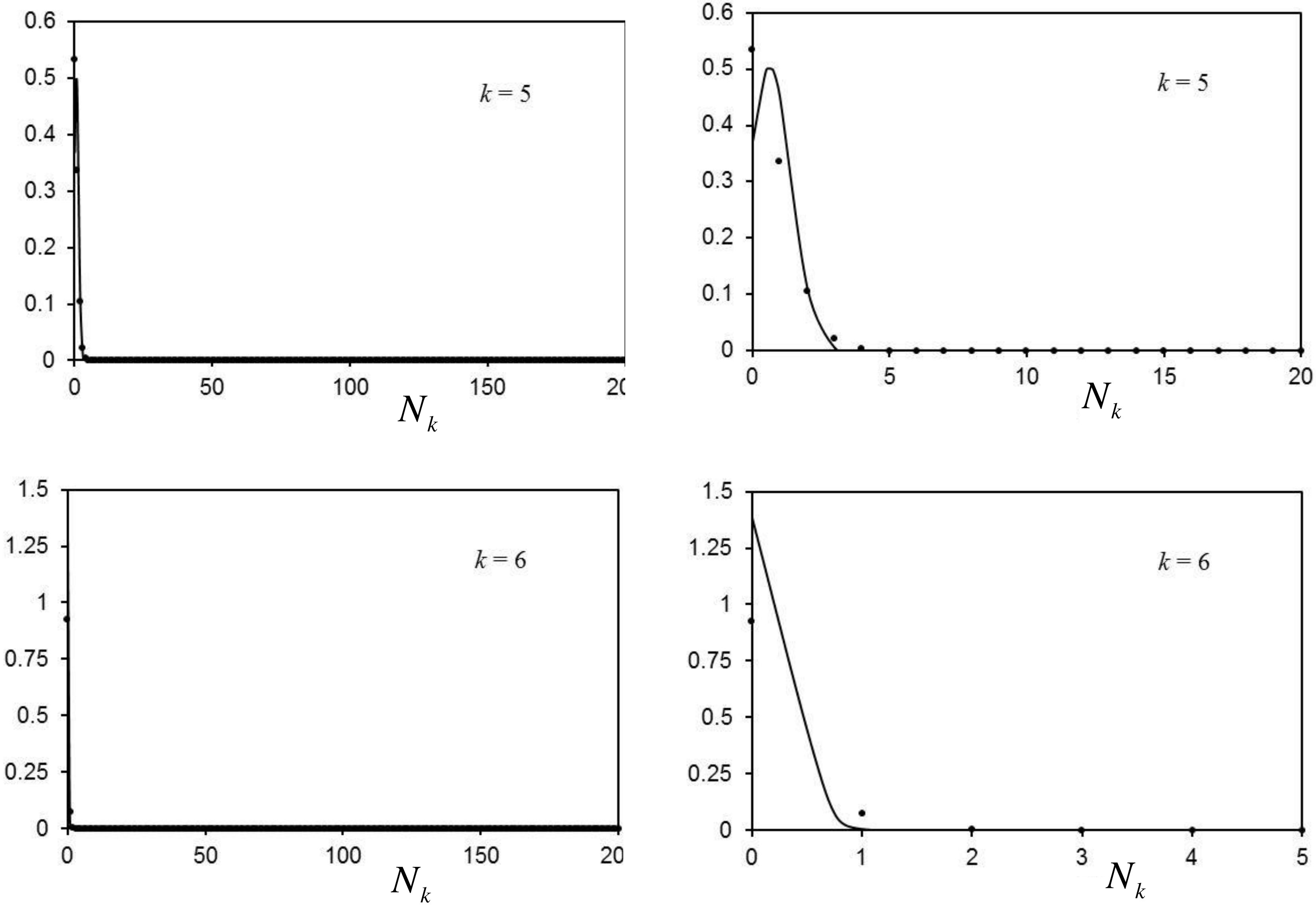
$$\sum_{k=1}^{K} \xi_k^{\ 2} \sim \chi^2_{K-2}$$

$$\sum_{k=1}^{K} N_k = 200, \quad \sum_{k=1}^{K} x_k N_k = 122, \quad K = 6$$

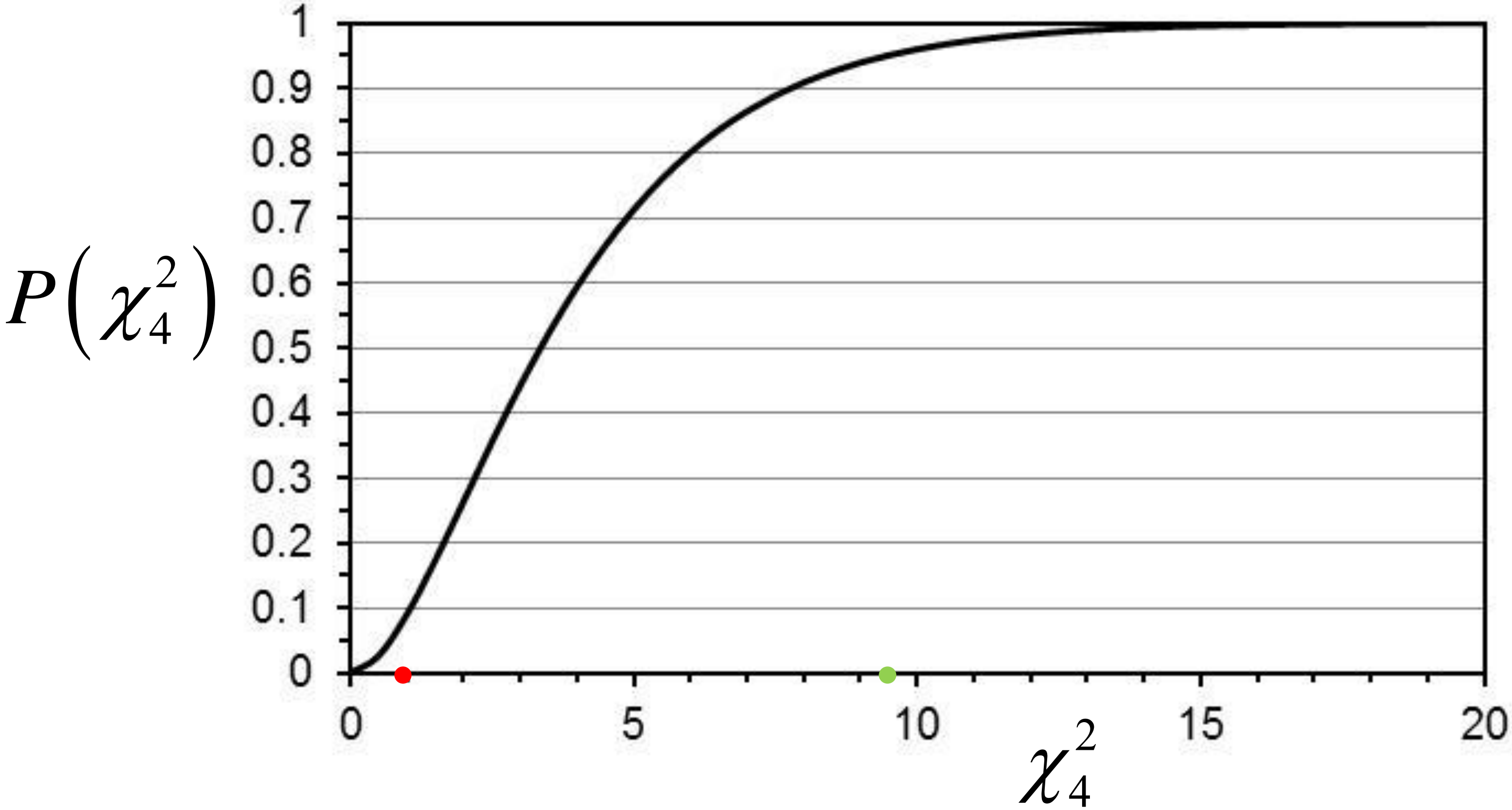
$$\sum_{k=1}^{K} \xi_k^{\ 2} = 0.820$$



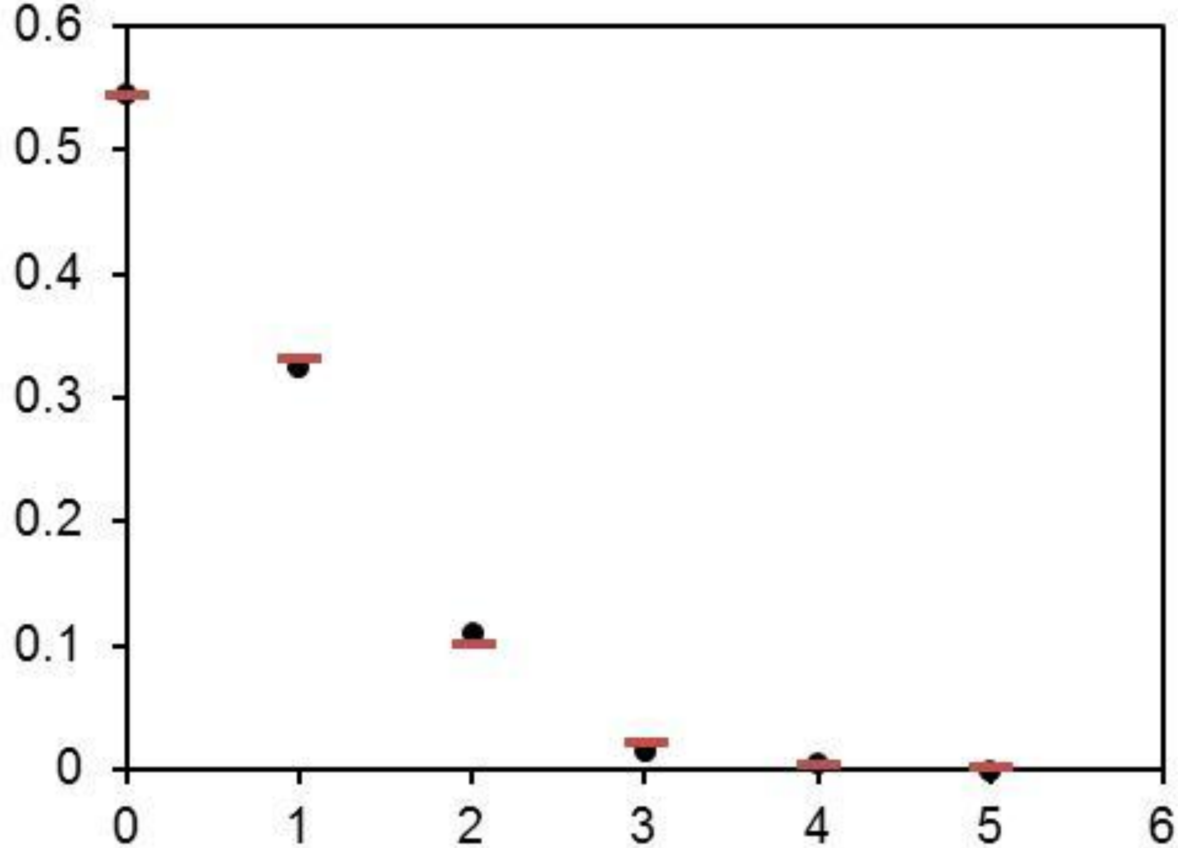
0.06
0.05
0.04
0.03
0.02
0.01
0
0
50
100
150
200
N_k
$k = 1$
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
50
100
150
200
N_k
$k = 2$
0.1
0.08
0.06
0.04
0.02
0
0
50
100
150
200
N_k
$k = 3$
0.25
0.2
0.15
0.1
0.05
0
0
50
100
150
200
N_k
$k = 4$



0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
$k = 5$
N_k
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
$k = 5$
N_k
1.5
1.25
1
0.75
0.5
0.25
0
0
50
100
150
$k = 6$
N_k
1.5
1.25
1
0.75
0.5
0.25
0
0
1
2
3
4
5
$k = 6$
N_k



$$P\left(\chi_4^2\right) = 0.95 \quad \text{при} \quad \chi_4^2 = 9.5$$



0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6