

Распределения выборочного среднего для норм. распр. и Коши

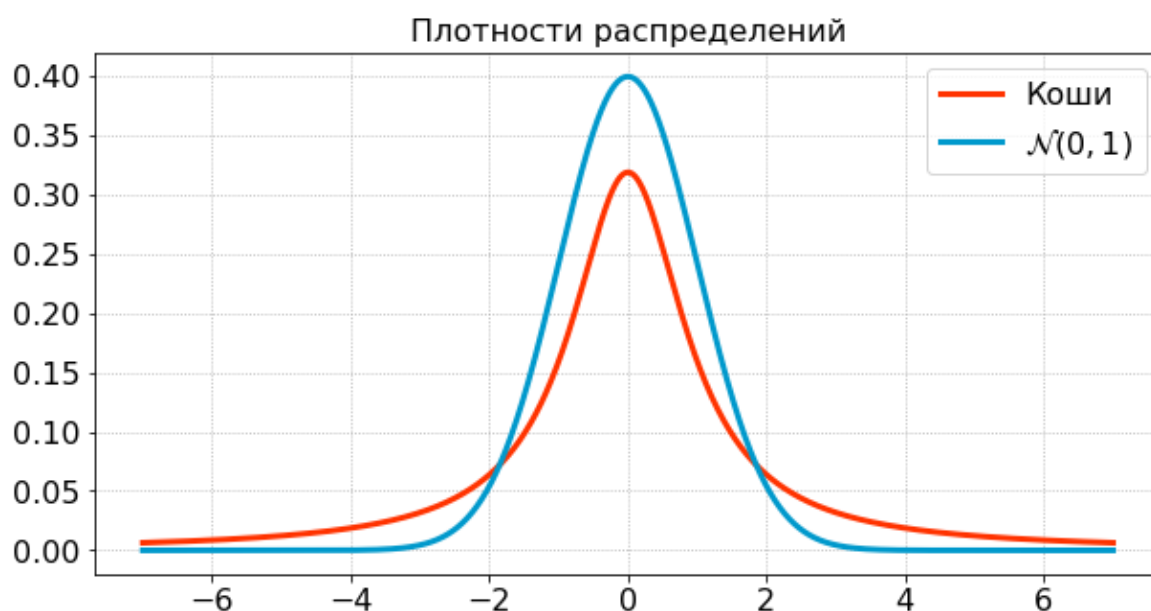
In [1]:

```
1 import numpy as np
2 import scipy.stats as sps
3 import matplotlib.pyplot as plt
4 %matplotlib inline
5
6 red = '#FF3300'
7 blue = '#0099CC'
8 green = '#00CC66'
```

Построим плотность стандартного нормального распределения и стандартного распределения Коши. Видим, что распределение Коши обладает более тяжелыми хвостами, нежели нормальное.

In [2]:

```
1 grid = np.linspace(-7, 7, 1000)
2
3 plt.figure(figsize=(10, 5))
4 plt.plot(grid, sps.cauchy.pdf(grid), color=red, label='Коши', lw=3)
5 plt.plot(grid, sps.norm.pdf(grid), color=blue,
6           label='$\mathcal{N}(0, 1)$', lw=3)
7 plt.xticks(fontsize=16), plt.yticks(fontsize=16)
8 plt.title('Плотности распределений', fontsize=16)
9 plt.legend(fontsize=16)
10 plt.grid(ls=':')
11 plt.show()
```



Сгенерируем 10^5 выборок размера 30 из стандартного нормального распределения

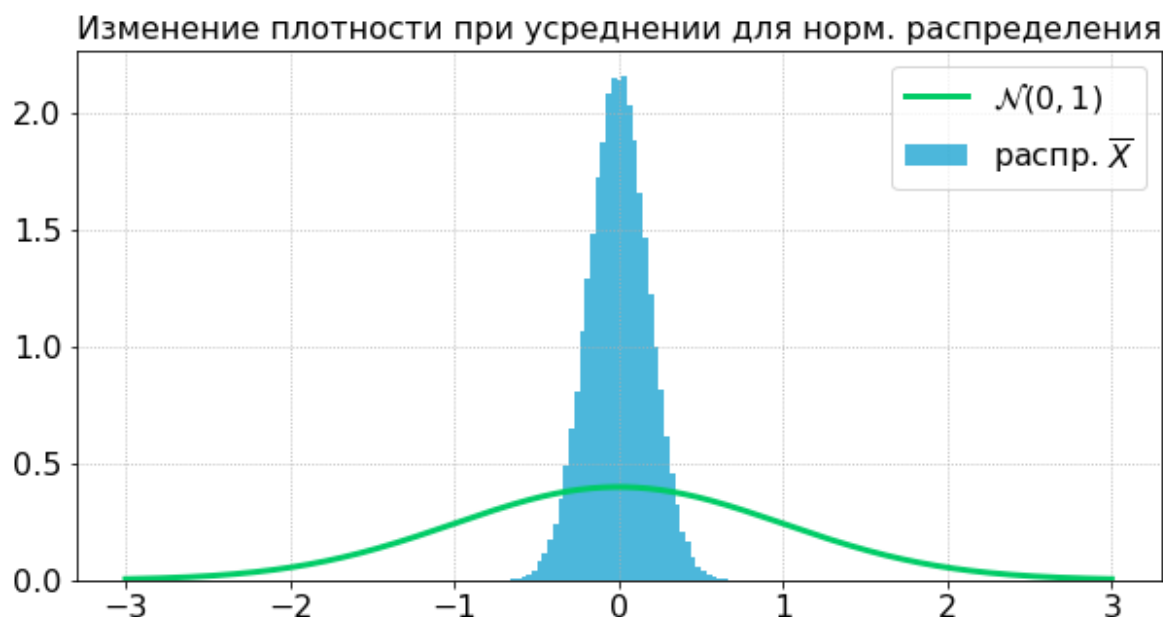
In [3]:

```
1 size = 30
2 count = 10 ** 5
3 sample = sps.norm.rvs(size=(count, size))
```

По каждой выборке из 30 элементов посчитаем выборочное среднее. Далее по ним построим гистограмму и построим плотность стандартного нормального распределения. Как видим, в подтверждении теории распределение выборочного среднего имеет меньшую дисперсию.

In [4]:

```
1 grid = np.linspace(-3, 3, 1000)
2
3 plt.figure(figsize=(10, 5))
4 plt.hist(sample.mean(axis=1), bins=50, density=True, alpha=0.7,
5          color=blue, label='распр.  $\overline{X}$ ')
6 plt.plot(grid, sps.norm.pdf(grid), color=green, lw=3,
7          label=' $\mathcal{N}(0, 1)$ ')
8 plt.xticks(fontsize=16), plt.yticks(fontsize=16)
9 plt.title('Изменение плотности при усреднении для норм. распределения',
10          fontsize=16)
11 plt.legend(fontsize=16)
12 plt.grid(ls=':')
13 plt.show()
```



Сгенерируем 10^5 выборок размера 30 из стандартного распределения Коши

In [5]:

```
1 size = 30
2 count = 10 ** 5
3 sample = sps.cauchy.rvs(size=(count, size))
```

Построим аналогичный график. В этом случае распределение выборочного среднего не отличается от распределения одного элемента выборки.

In [6]:

```
1 grid = np.linspace(-10, 10, 1000)
2
3 plt.figure(figsize=(10, 5))
4 plt.hist(sample.mean(axis=1), bins=50, density=True, alpha=0.7,
5          color=green, label='распр.  $\overline{X}$ ', range=(-10, 10))
6 plt.plot(grid, sps.cauchy.pdf(grid), color=red, lw=3,
7          label='Коши')
8 plt.xticks(fontsize=16), plt.yticks(fontsize=16)
9 plt.title('Изменение плотности при усреднении для распределения Коши',
10          fontsize=16)
11 plt.legend(fontsize=16)
12 plt.grid(ls=':')
13 plt.show()
```



Статистика, прикладной поток 2019

<https://mipt-stats.gitlab.io/> (<https://mipt-stats.gitlab.io/>)