Задание 6.3

In [3]:

```
%matplotlib inline
import scipy.stats as st
import numpy as np
import matplotlib.pyplot as plt
import math
import statsmodels.api as sm
```

Сгенерируем выборку Коши размера $N = 100$

In [4]:

```
N = 100
sample = st.cauchy.rvs(size=100)
```

Посчитаем ОМП распределения Коши, как модель $N(\theta, 1)$ (т.е. выборочное среднее)

In [5]:

```
thetas = np.zeros(N+1)
for i in range(N+1):
        thetas[i] = np.mean(sample[:(i+1)])
```

Для нормального распределения $N(\theta, 1)$ априорное распределение $N(\theta_0, \sigma_0^2)$ с мат ожиданием равным θ_0. Байесовской оценкой параметра будет: $\hat{\theta} = \frac{\frac{\sum_{i=1}^n X_i}{\sigma^2} + \frac{\theta_0}{\sigma_0^2}}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}}$, где $\sigma^2 = 1$

Функция байесовской оценки

$\hat{\theta} = \frac{\sum_{i=1}^n X_i}{\frac{1}{\sigma^2} + n}$.

In [6]:

```
def Bayesian(theta, sigma_0 , x):
    a = sum(x) + (theta/sigma_0)
    b = (1/sigma_0) + len(x)
    return (a/b)
```

Теперь требуется по условию задачи так подобрать параметры априорного распределения, что с вероятностью не менее 0.95 выполнено неравенство $|\theta| < 0.5$. То есть, $P(-0.5 < \theta < 0.5) \geq 0.95$

- $\theta_0 = 0$
- Найдем параметр σ_0.
- Применим неравенство Чебышева:
- $P(|\theta| > 0.5) \leq \frac{D\theta}{0.5^2} = \frac{\sigma_0^2}{0.25}$

- $P(|\theta| < 0.5) = 1 - P(|\theta| > 0.5) \geq 1 - \frac{\sigma_0^2}{0.25}$
- $\frac{\sigma_0^2}{0.25} \geq 1 - P(|\theta| < 0.5) = 0.05$
- Возьмем $\sigma^2 = 0.0125.$

Тогда параметры априорного распределения: $\theta_0 = 0; \sigma_0 = 0.25$

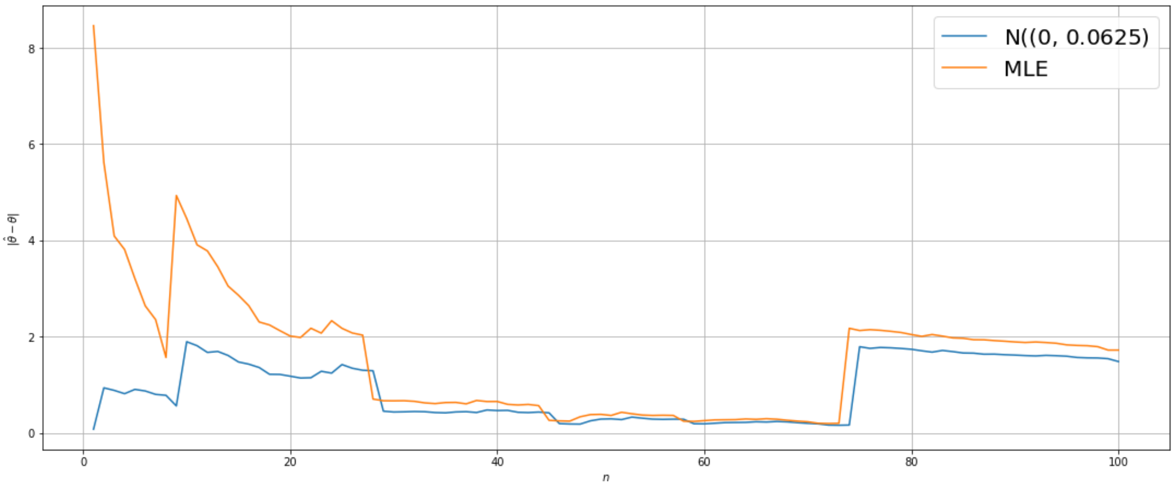
In [7]:

```
params = ([0,1/16])
fig = plt.gcf()
plt.grid(True)
fig.set_size_inches(18.5, 7.5)
stat_bayes = np.zeros(N+1)
alpha_0 = params[0]
beta_0 = params[1]
for j in range(1, N+1):
    stat_bayes[j] = Bayesian(alpha_0, beta_0, sample[:j])
plt.plot(np.arange(1,N+1), abs(stat_bayes[1:] - 0), label='N$( \alpha, \beta)=$({}, {})'.fc

plt.plot(np.arange(1,N+1), abs(thetas[1:] - 0), label = r'MLE')
plt.xlabel(r'$n$')
plt.ylabel(r'$|\hat{\theta} - \theta |$')
plt.legend(fontsize=20, loc=1)
```

Out[7]:

```
<matplotlib.legend.Legend at 0x1e39596ae10>
```



Вывод

Так как на самом деле выборка из распределения Коши, которое не имеет среднего, и мы рассматривали ее как модель $N(\theta, 1)$, то график выше показывает, что байесовская оценка ведет себя лучше (в смысле абсолютного отклонения от истинного параметра), чем ОМП. И связано это еще с тем, что мы удачно подобрали параметры. То есть байесовская оценка уменьшает влияние выборки на оценку параметра. Чем меньше сигма в априорном распределении, тем больше априорное знание влияет на оценку(с большим весом). Грубо говоря можно подогнать оценку параметра с помощью байесовской оценке.

In [ ]:

```

```