Задача 4.1

In [2]:

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

Сгенерируем по 100 выборок для каждого параметра θ

In [3]:

```
M = 100
size_s = 1000
tetas = np.array([1,5,10])
sample =np.zeros((3, M, size_s))
for j in range(tetas.size):
    for i in range(M):
        uniform = st.uniform(loc=0, scale=tetas[j])
        sample[j][i] = uniform.rvs(size=size_s)
```

1) Считаем оценку $2\overline{X}$ и найдем среднее значение функции потерь для фиксированного n.

In [4]:

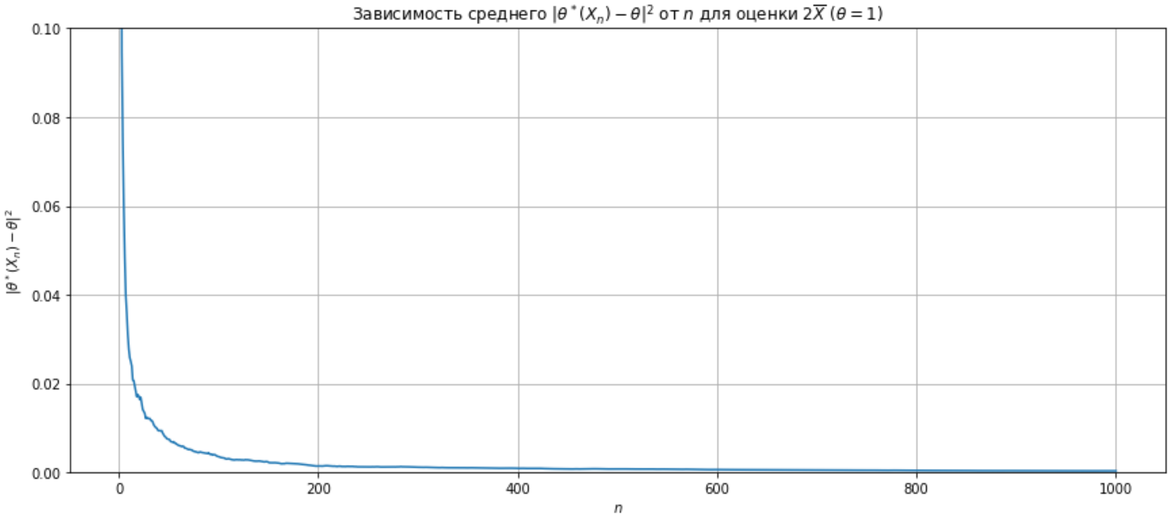
```
stat1 = np.zeros((3, M, size_s+1))
less_function1 = np.zeros((3, M, size_s+1))
mean_less_func1 = np.zeros((3, size_s+1))
for j in range(tetas.size):
    for i in range(M):
        for q in range(1, size_s+1):
            stat1[j][i][q] = 2*np.mean(sample[j][i][:q])
            less_function1[j][i][q] = (stat1[j][i][q] - tetas[j])**2

for j in range(tetas.size):
    for q in range(1, size_s+1):
        mean_less_func1[j][q] = np.average(less_function1[j,:,q])

print (mean_less_func1[:,1000])
```

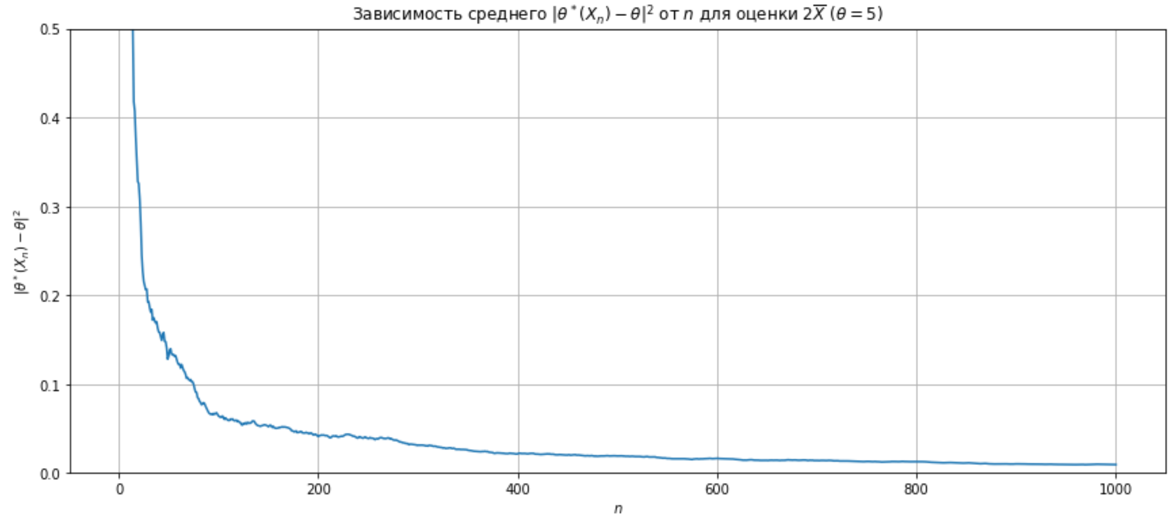
```
[0.00034024 0.00975535 0.03370134]
```

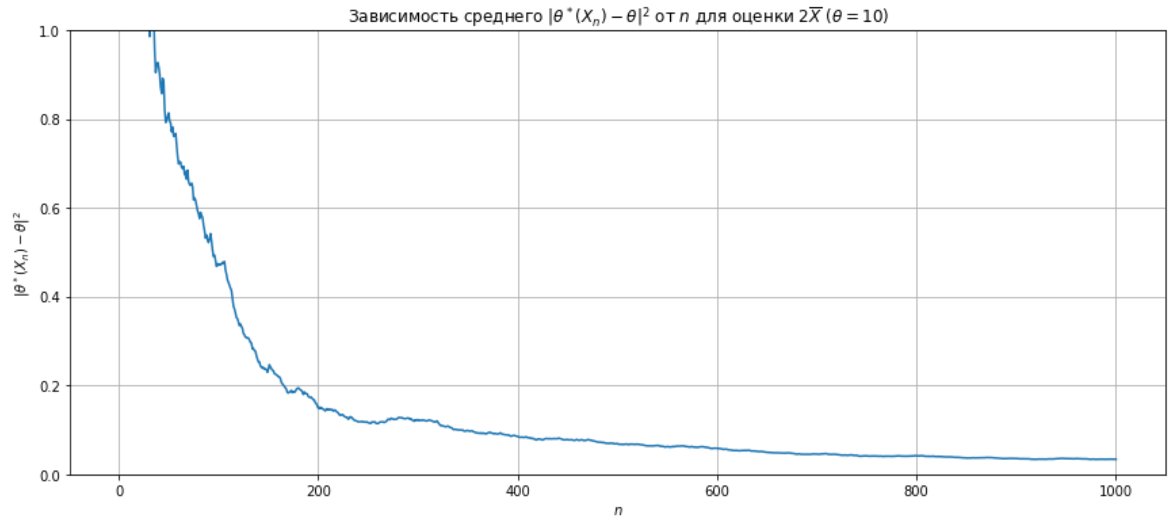
Для себя выведем графики зависимости функции потерь для оценки $2\overline{X}$

In [5]:

```
for i in range(tetas.size):
    plt.subplot (3, 1, i+1)
    plt.plot(np.arange(1,1001), mean_less_func1[i][1:])
    plt.xlabel(r'$n$')
    plt.ylabel(r'$|\theta^*(X_{n}) - \theta |^2$')
    plt.title(r'Зависимость среднего $|\theta^*(X_{n}) - \theta |^2$ от $n$ для оценки $2\c
    plt.grid(True)
    plt.ylim (0, 0.10*tetas[i])
    fig = plt.gcf()
    fig.set_size_inches(14.5, 20.5)
```







С увеличением параметра θ, график зависимости среднего квадратичного поднимаетя выше.

2) Считаем оценку $(n+1)X_{(1)}$ и найдем среднее значение функции потерь для фиксированного n.

In [6]:

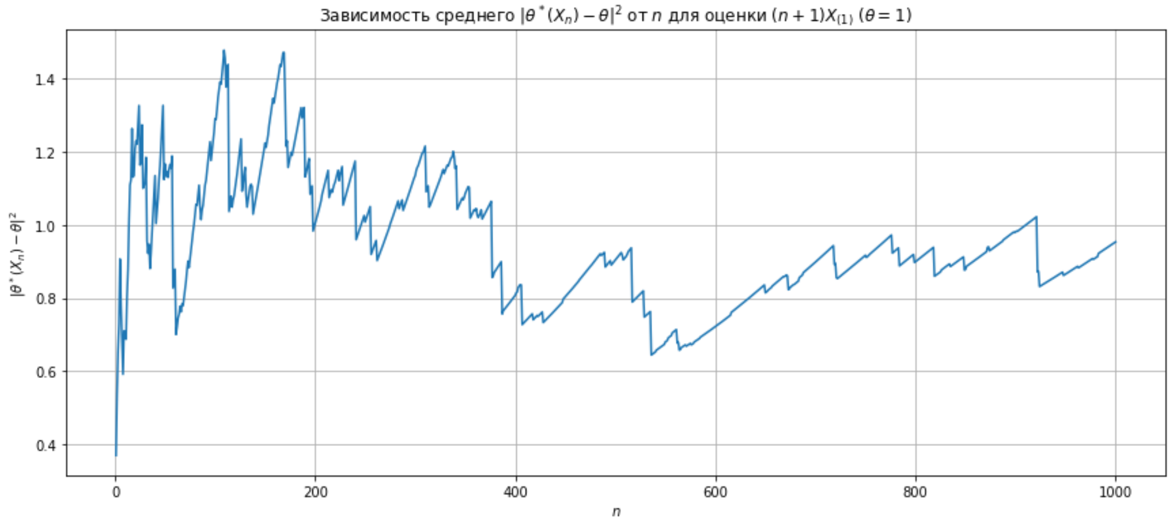
```
stat2 = np.zeros((3, M, size_s+1))
less_function2 = np.zeros((3, M, size_s+1))
mean_less_func2 = np.zeros((3, size_s+1))
for j in range(tetas.size):
    for i in range(M):
        for q in range(1, size_s+1):
            stat2[j][i][q] = (q + 1.0)*np.amin(sample[j][i][:q])
            less_function2[j][i][q] = (stat2[j][i][q] - tetas[j])**2

for j in range(tetas.size):
    for q in range(1, size_s+1):
        mean_less_func2[j][q] = np.average(less_function2[j,:,q])

print (mean_less_func2[:,1000])
```

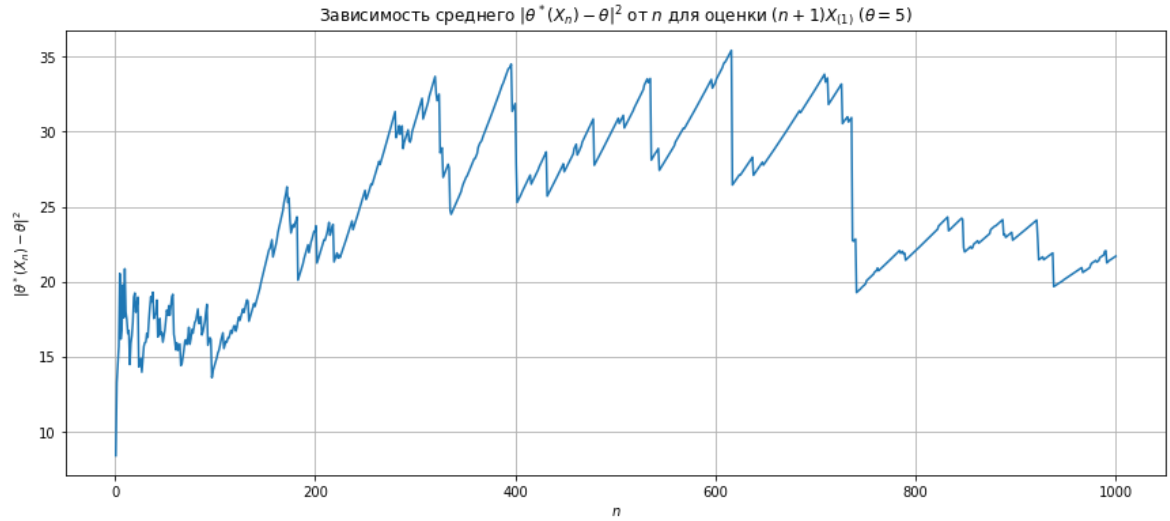
```
[ 0.95372306 21.69198357 86.24210517]
```

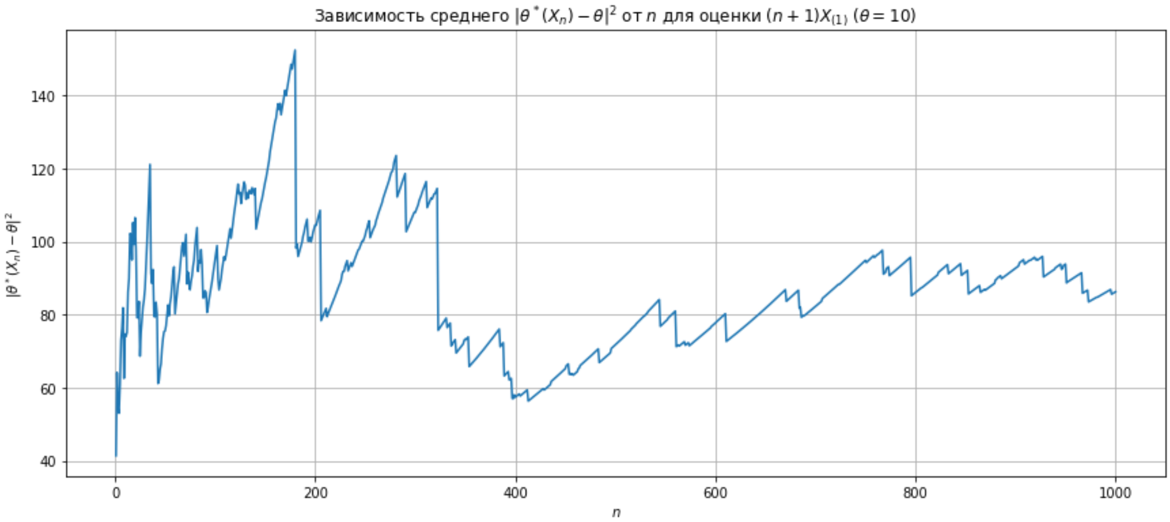
Для себя построим график

In [7]:

```
for i in range(tetas.size):
    plt.subplot (3, 1, i+1)
    plt.plot(np.arange(1,1001), mean_less_func2[i][1:])
    plt.xlabel(r'$n$')
    plt.ylabel(r'$|\theta^*(X_{n}) - \theta |^2$')
    plt.title(r'Зависимость среднего $|\theta^*(X_{n}) - \theta |^2$ от $n$ для оценки $(n+
    plt.grid(True)
    #plt.ylim (0, 0.10*tetas[i])
    fig = plt.gcf()
    fig.set_size_inches(14.5, 20.5)
```







Из графиков видно, что оценку нельзя назвать состоятельной и сравнии с другими оценками она очень сильно отличается от теоретического значения параметра, следовательно далее, уберем из рассмотрения эту оценку т.к. она не состоятельна.

3) Считаем оценку $X_{(1)} + X_{(n)}$ и найдем среднее значение функции потерь для фиксированного n.

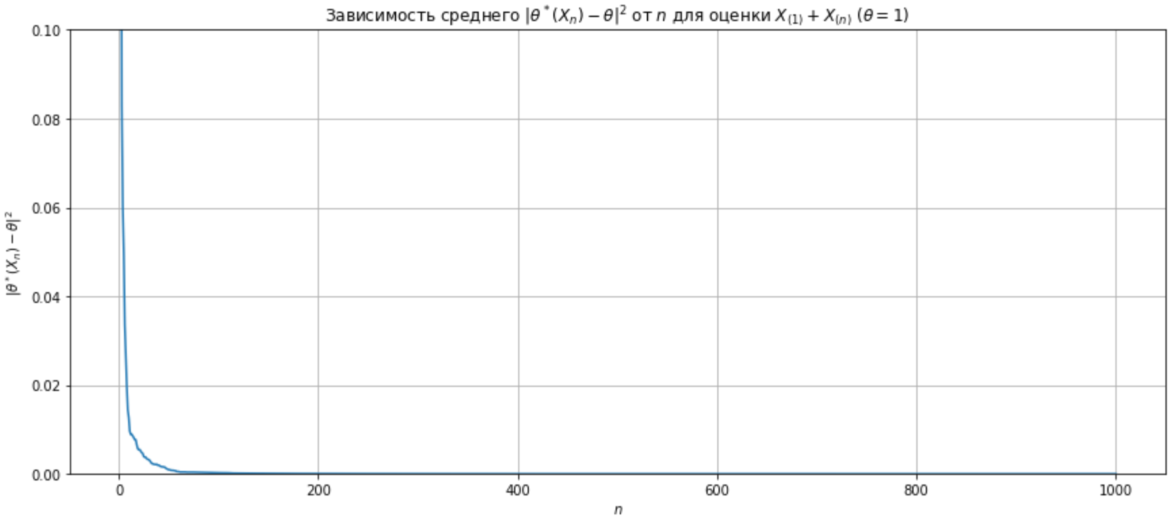
In [8]:

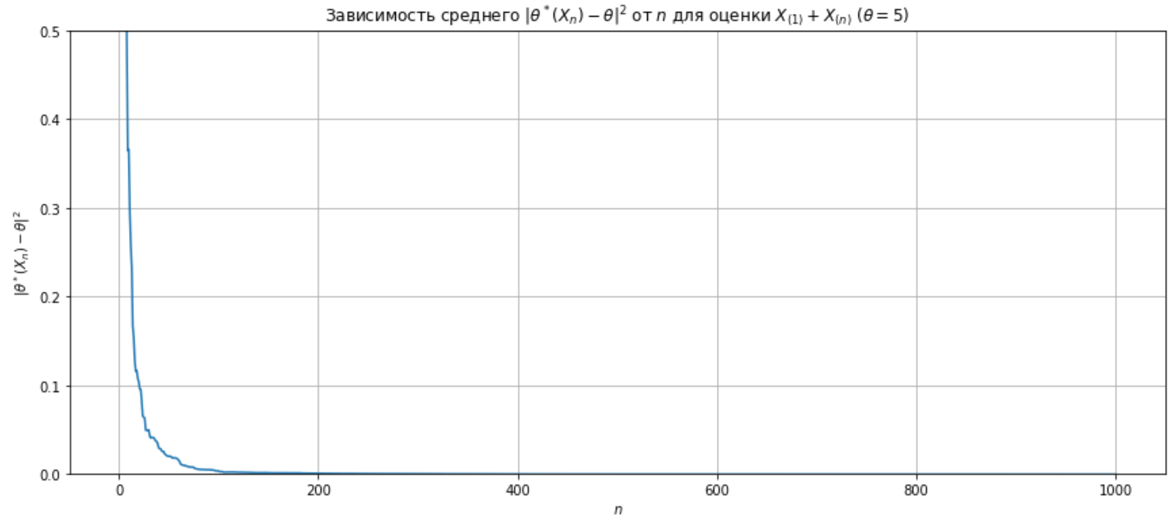
```
stat3 = np.zeros((3, M, size_s+1))
less_function3 = np.zeros((3, M, size_s+1))
mean_less_func3 = np.zeros((3, size_s+1))
for j in range(tetas.size):
    for i in range(M):
        for q in range(1, size_s+1):
            stat3[j][i][q] = np.amin(sample[j][i][:q]) + np.amax(sample[j][i][:q])
            less_function3[j][i][q] = (stat3[j][i][q] - tetas[j])**2

for j in range(tetas.size):
    for q in range(1, size_s+1):
        mean_less_func3[j][q] = np.average(less_function3[j,:,q])

print (mean_less_func3[:,1000])
```

```
[1.63666564e-06 4.14898832e-05 1.53089599e-04]
```

In [9]:

```
for i in range(tetas.size):
    plt.subplot (3, 1, i+1)
    plt.plot(np.arange(1,1001), mean_less_func3[i][1:])
    plt.xlabel(r'$n$')
    plt.ylabel(r'$|\theta^*(X_{n}) - \theta |^2$')
    plt.title(r'Зависимость среднего $|\theta^*(X_{n}) - \theta |^2$ от $n$ для оценки $X_{
    plt.grid(True)
    plt.ylim (0, 0.10*tetas[i])
    fig = plt.gcf()
    fig.set_size_inches(14.5, 20.5)
```





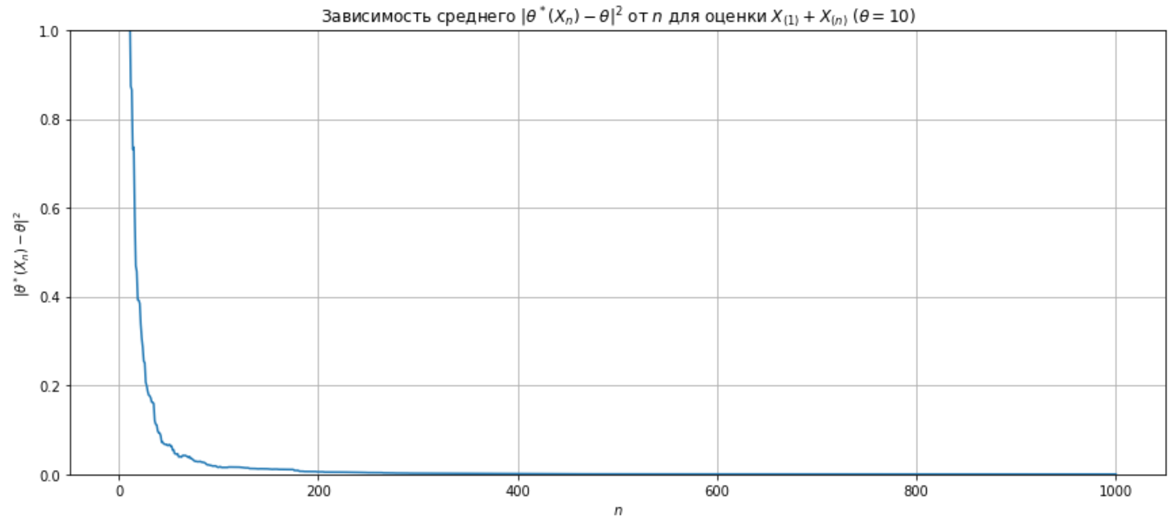


График становится выше с ростом параметра

4) Считаем оценку $\frac{n+1}{n}X_{(n)}$ и найдем среднее значение функции потерь для фиксированного n.

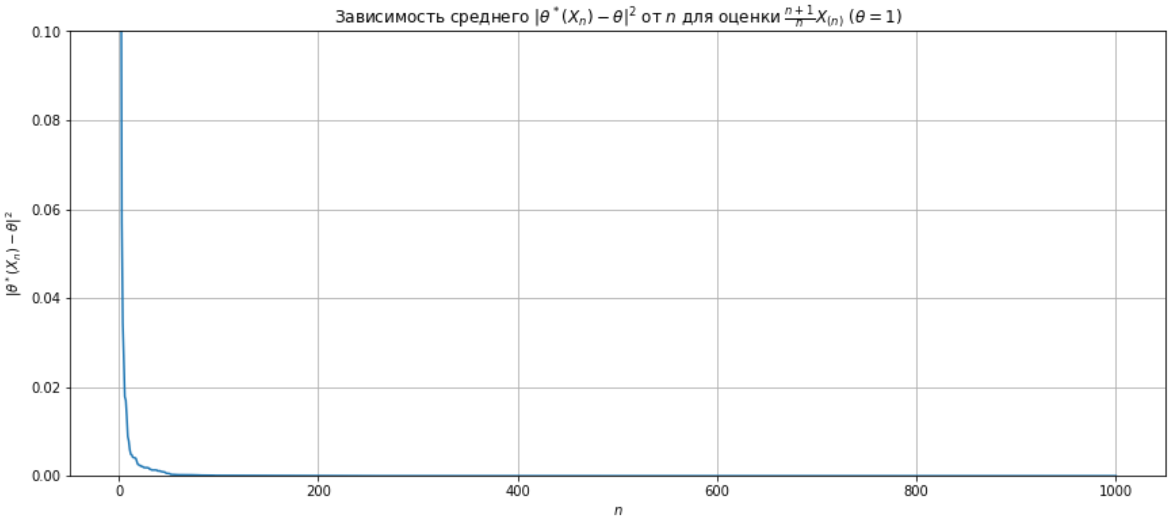
In [10]:

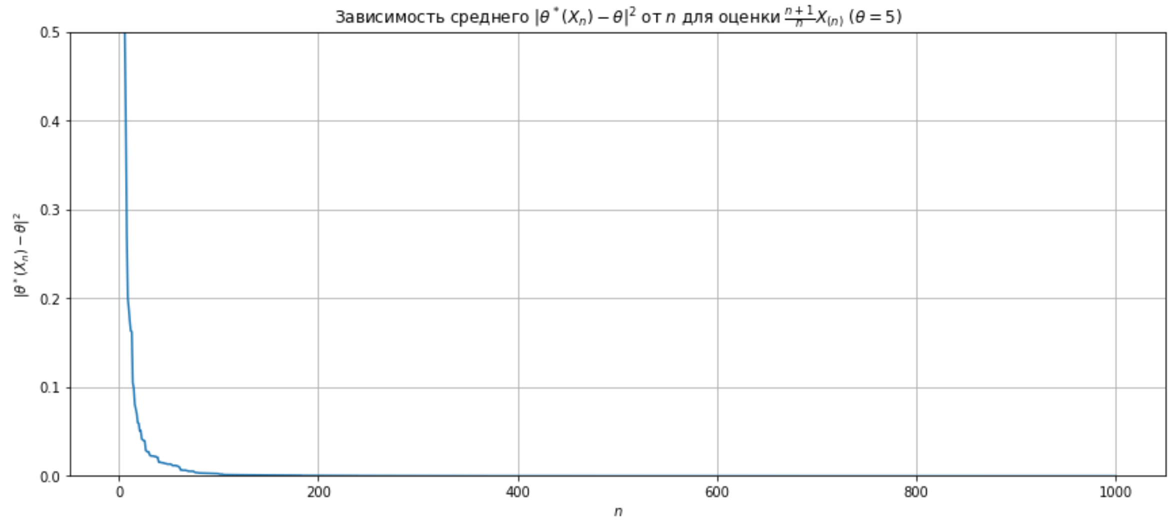
```
stat4 = np.zeros((3, M, size_s+1))
less_function4 = np.zeros((3, M, size_s+1))
mean_less_func4 = np.zeros((3, size_s+1))
for j in range(tetas.size):
    for i in range(M):
        for q in range(1, size_s+1):
            stat4[j][i][q] =  np.amax(sample[j][i][:q]) * (q+1.0)/q
            less_function4[j][i][q] = (stat4[j][i][q] - tetas[j])**2

for j in range(tetas.size):
    for q in range(1, size_s+1):
        mean_less_func4[j][q] = np.average(less_function4[j,:,q])

print (mean_less_func4[:,1000])
```

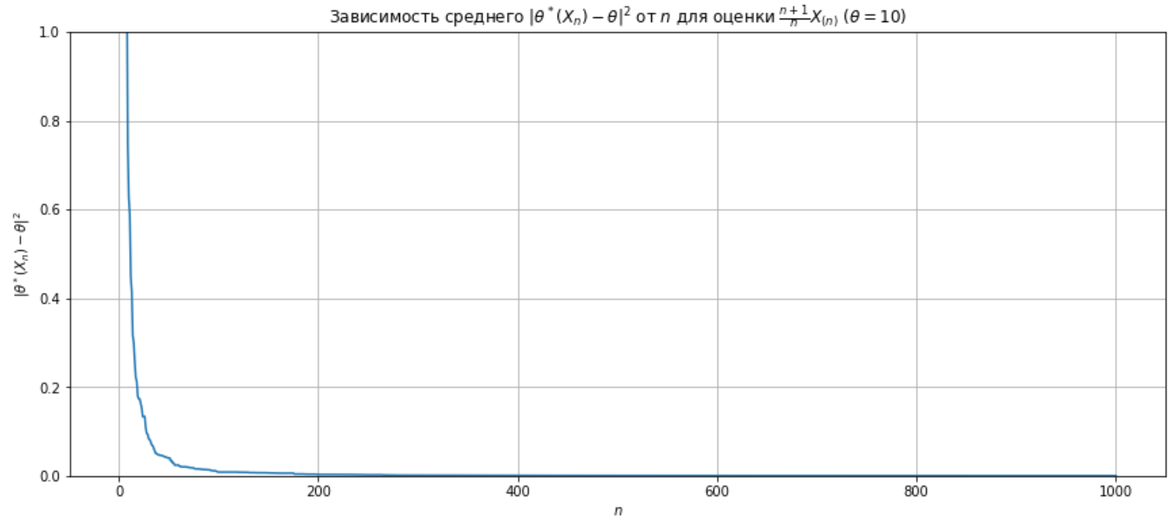
```
[6.48476413e-07 2.18321759e-05 7.75014572e-05]
```

In [11]:

```
for i in range(tetas.size):
    plt.subplot (3, 1, i+1)
    plt.plot(np.arange(1,1001), mean_less_func4[i][1:])
    plt.xlabel(r'$n$')
    plt.ylabel(r'$|\theta^*(X_{n}) - \theta |^2$')
    plt.title(r'Зависимость среднего $|\theta^*(X_{n}) - \theta |^2$ от $n$ для оценки $\fr
    plt.grid(True)
    plt.ylim (0, 0.10*tetas[i])
    fig = plt.gcf()
    fig.set_size_inches(14.5, 20.5)
```



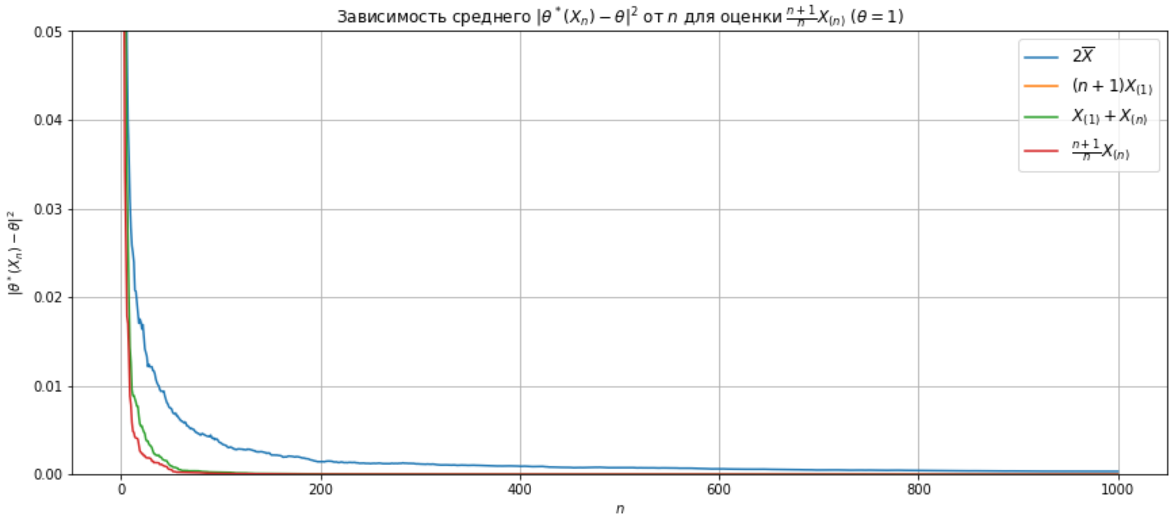


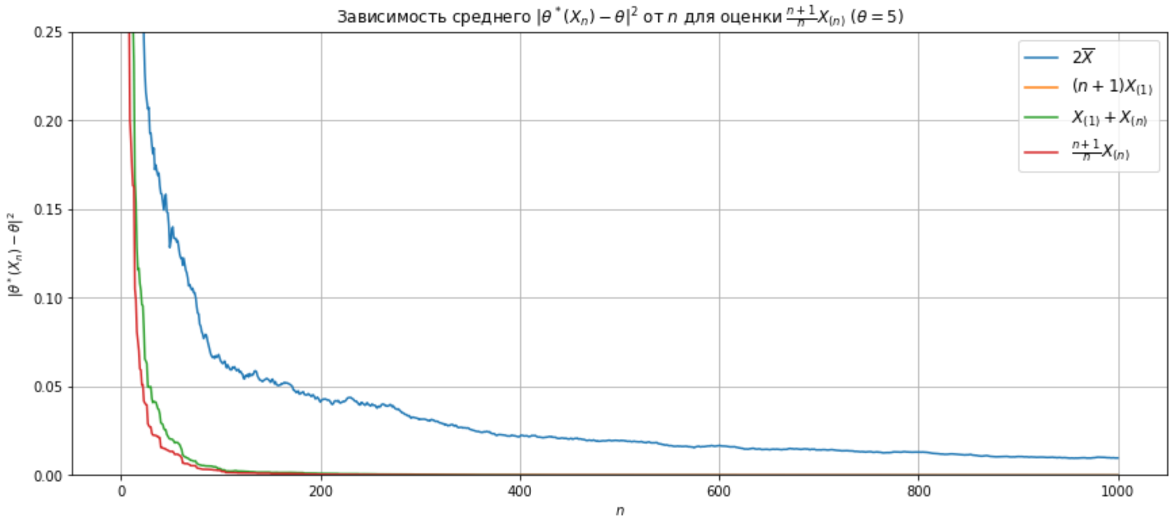


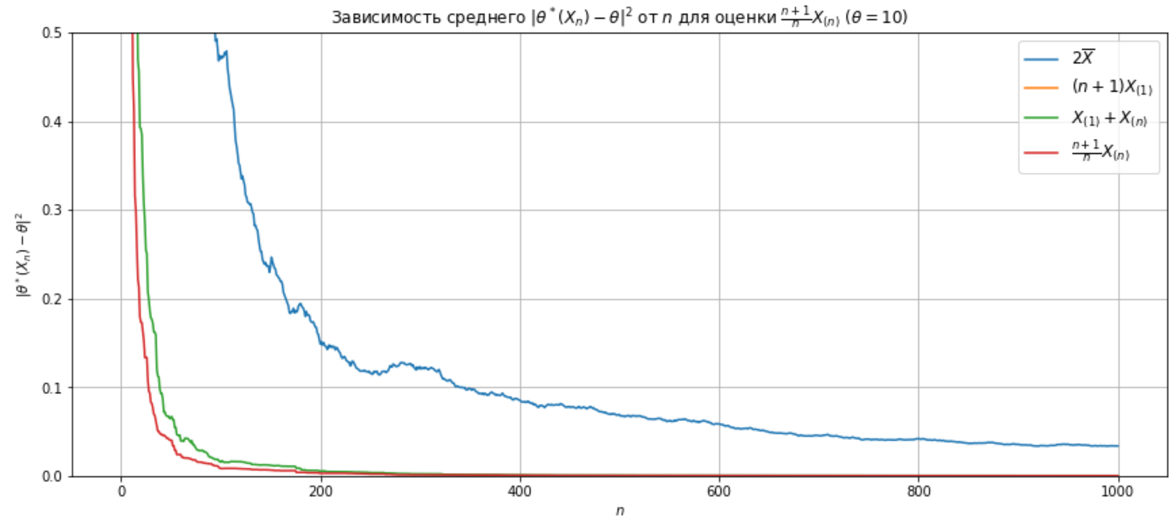
Проанализируем оценки на одном графике для фиксированного θ

In [13]:

```
for i in range(tetas.size):
    plt.subplot(3, 1, i+1)
    plt.plot(np.arange(1,1001), mean_less_func1[i][1:], label = r'$2\overline{X}$')
    #plt.plot(np.arange(1,1001), mean_less_func2[i][1:], label = r'$(n+1)X_{(1)}$') # Несос
    plt.plot(np.arange(1,1001), mean_less_func3[i][1:], label = r'$X_{(1)} + X_{(n)}$')
    plt.plot(np.arange(1,1001), mean_less_func4[i][1:], label = r'$\frac{n+1}{n}X_{(n)}$')
    plt.xlabel(r'$n$')
    plt.ylabel(r'$|\theta^*(X_{n}) - \theta |^2$')
    plt.legend(fontsize=12, loc=1)
    plt.title(r'Зависимость среднего $|\theta^*(X_{n}) - \theta |^2$ от $n$ для оценки $\fr
    plt.grid(True)
    plt.ylim (0, 0.05*tetas[i])
    fig = plt.gcf()
    fig.set_size_inches(14.5, 20.5)
```







Вывод

Было доказано(на семинаре), что для выборки из равомерного закона $U[0,\theta]$ сравнимы следующие оценки параметра θ в равномерном подходе с квадратичной функцией потерь. Причем получено, что оценка $\frac{n+1}{n}X_{(n)}$ лучше оценки $2\overline{X}$. Что можно наблюдать на построенных графиков для трех значений θ. Также исходя из графиков, можно сказать, что в равномерном подходе с квадратичной функцией потерь оценка $X_{(1)}+X_{(n)}$ хуже чем оценка $\frac{n+1}{n}X_{(n)}$, но лучше чем $2\overline{X}$. Хуже всех повела себя оценка $(n+1)X_{(1)}$, так как она не является состоятельной.