

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as sts
from statsmodels.distributions.empirical_distribution import ECDF
%matplotlib inline
```

Выборка размера $N = 10000$ из $N(0, 1)$

```
In [3]: N = 10000
mu = 2.0
sigma = 0.5

norm_rv = sts.norm(loc=mu, scale=sigma)

# сгенерируем N значений
sample = norm_rv.rvs(N)
```

Для каждого $n \leq N$ посчитаем его эмпирическую функцию распределения

```
In [5]: ecdf = np.array([])
for x in xrange(1, N + 1):
    ecdf = np.append(ecdf, ECDF(sample[:x]))
```

График эмпирических функций распределения для $n = \{10, 25, 50, 100, 1000, N\}$

```
In [25]: x = np.linspace(0,4,100)

plt.step(ecdf[9].x, ecdf[9].y, label='$n = 10$')
plt.step(ecdf[24].x, ecdf[24].y, label='$n = 25$')
plt.step(ecdf[49].x, ecdf[49].y, label='$n = 50$')
plt.step(ecdf[99].x, ecdf[99].y, label='$n = 100$')
plt.step(ecdf[999].x, ecdf[999].y, label='$n = 1000$')
plt.step(ecdf[9999].x, ecdf[9999].y, label='$n = N$')
plt.plot(x, norm_rv.cdf(x), label='$F(x)$')

plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.show()
```

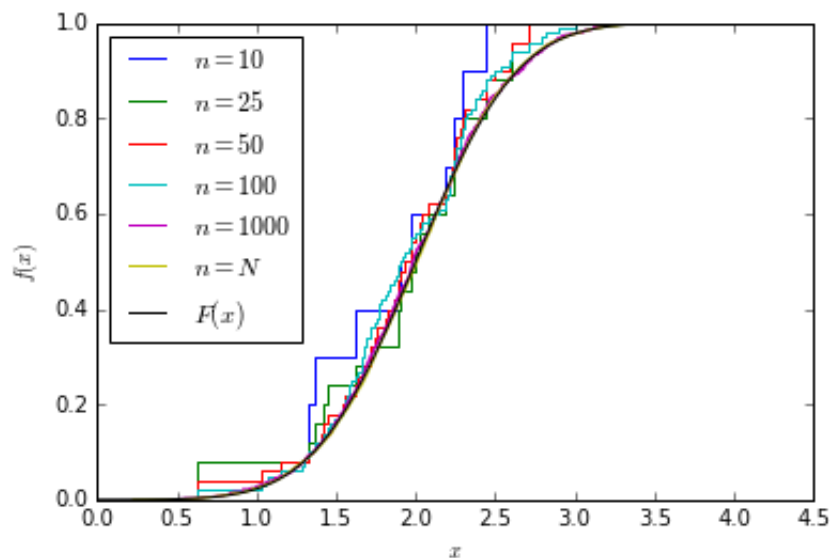
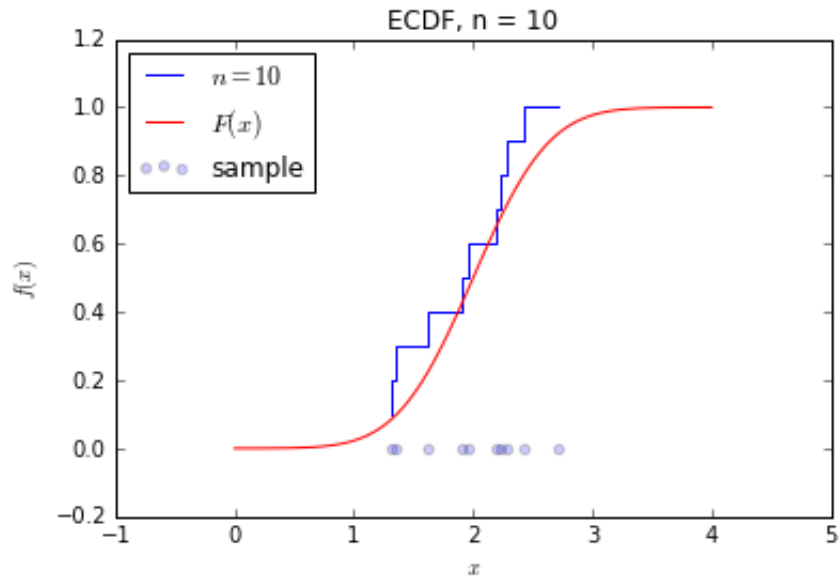


График эмпирической функций распределения для $n = 10$

```
In [26]: plt.step(ecdf[9].x, ecdf[9].y, label='$n = 10$')
plt.scatter(sample[:10], sample[:10] * 0, alpha=0.2, s=20, label='sample')
plt.plot(x, norm_rv.cdf(x), label='$F(x)$', color = 'red')

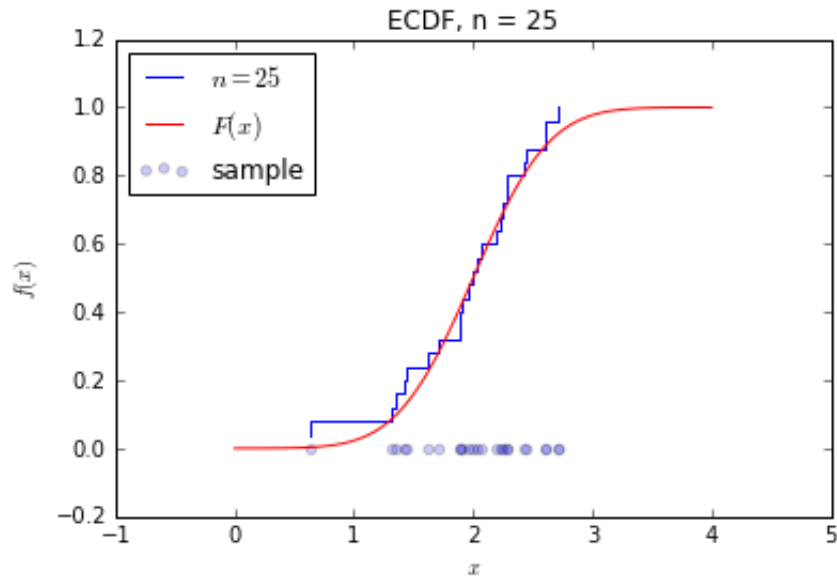
plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.title("ECDF, n = 10")
plt.show()
```



**График эмпирической функций распределения
для $n = 25$**

```
In [27]: plt.step(ecdf[24].x, ecdf[24].y, label='$n = 25$')
plt.scatter(sample[:25], sample[:25] * 0, alpha=0.2, s=20, label='sample')
plt.plot(x, norm_rv.cdf(x), label='$F(x)$', color = 'red')

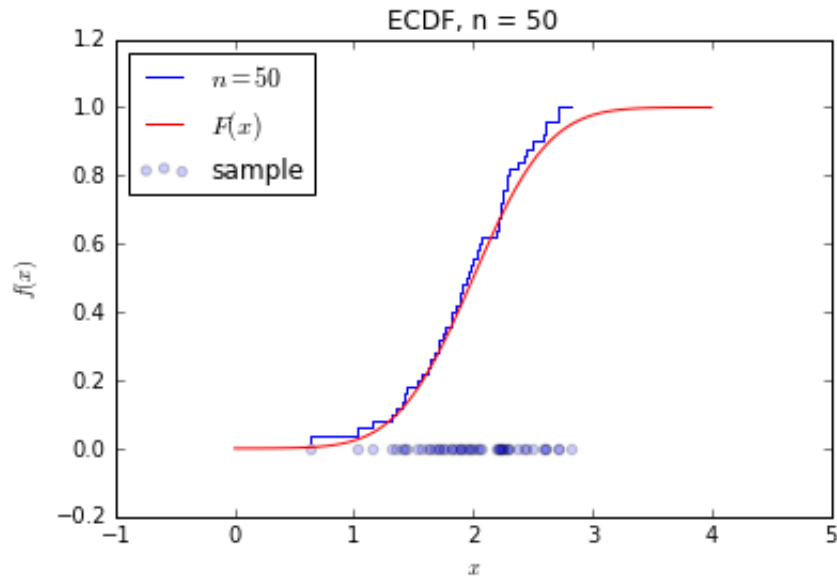
plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.title("ECDF, n = 25")
plt.show()
```



**График эмпирической функций распределения
для $n = 50$**

```
In [28]: plt.step(ecdf[49].x, ecdf[49].y, label='$n = 50$')
plt.scatter(sample[:50], sample[:50] * 0, alpha=0.2, s=20, label='sample')
plt.plot(x, norm_rv.cdf(x), label='$F(x)$', color = 'red')

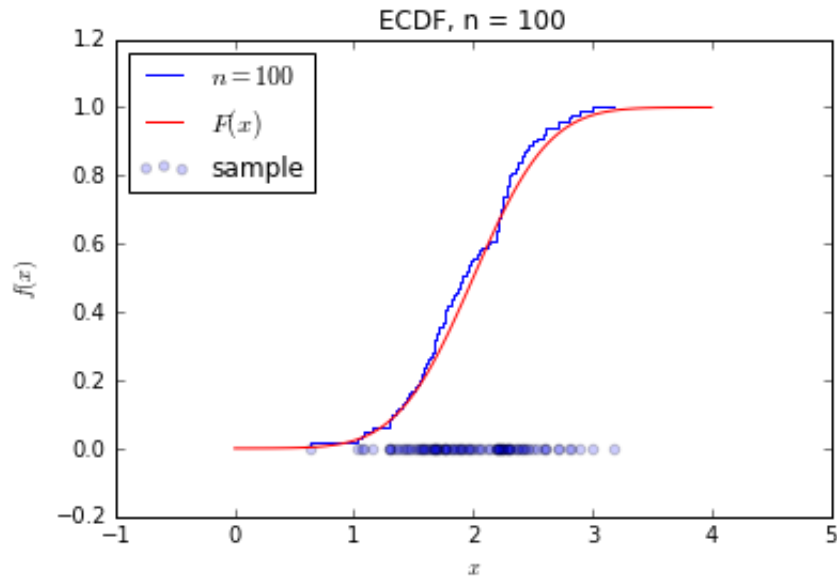
plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.title("ECDF, n = 50")
plt.show()
```



**График эмпирической функций распределения
для $n = 100$**

```
In [29]: plt.step(ecdf[99].x, ecdf[99].y, label='$n = 100$')
plt.scatter(sample[:100], sample[:100] * 0, alpha=0.2, s=20, label='sam
plt.plot(x, norm_rv.cdf(x), label='$F(x)$', color = 'red')

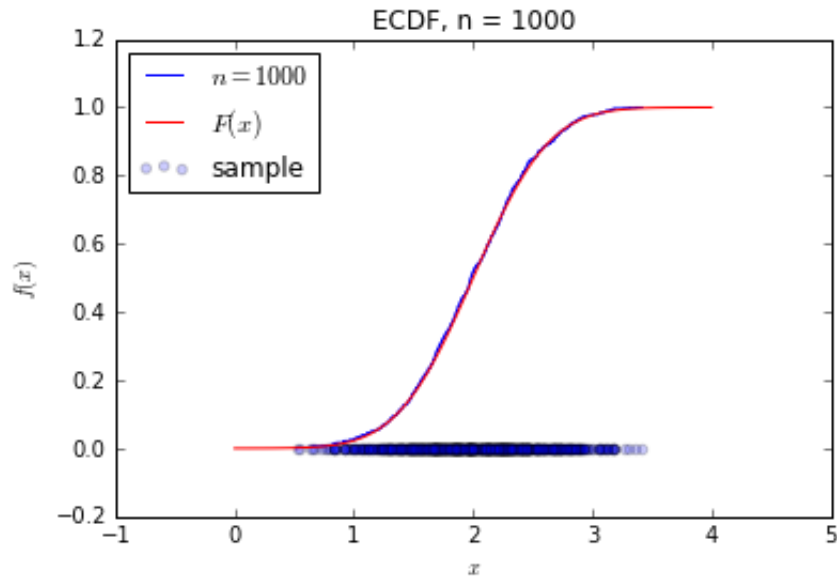
plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.title("ECDF, n = 100")
plt.show()
```



**График эмпирической функций распределения
для $n = 1000$**

```
In [30]: plt.step(ecdf[999].x, ecdf[999].y, label='$n = 1000$')
plt.scatter(sample[:1000], sample[:1000] * 0 ,alpha=0.2, s=20, label='s')
plt.plot(x, norm_rv.cdf(x), label='$F(x)$', color = 'red')

plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.title("ECDF, n = 1000")
plt.show()
```



**График эмпирической функций распределения
для $n = N$**

```
In [31]: plt.step(ecdf[N-1].x, ecdf[N-1].y, label='$n = N$')
plt.scatter(sample[:N], sample[:N] * 0, alpha=0.2, s=20, label='sample')
plt.plot(x, norm_rv.cdf(x), label='$F(x)$', color = 'red')

plt.ylabel('$f(x)$')
plt.xlabel('$x$')
plt.legend(loc='upper left')
plt.title("ECDF, n = N")
plt.show()
```

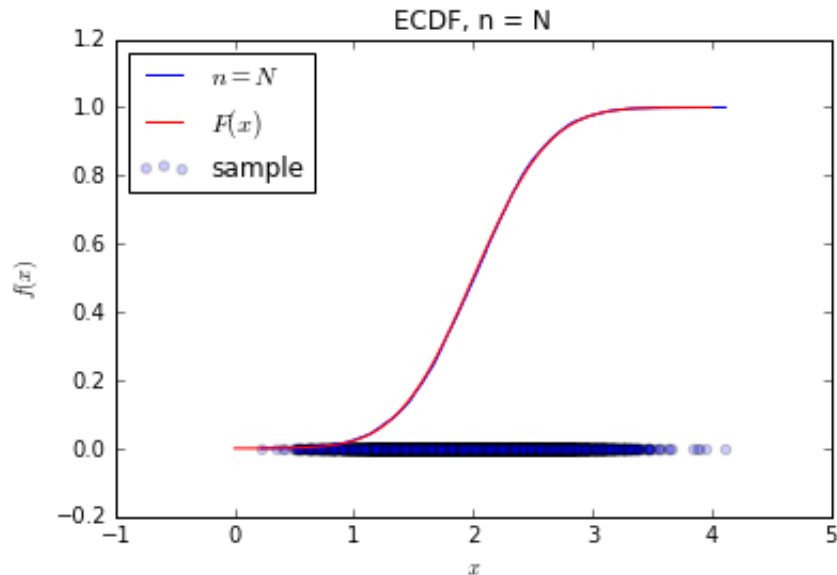


График $D_n = \sup_{x \in R} |F_n(x) - F(x)|$

```
In [43]: D = np.array([])
for x in xrange(0, N):
    D = np.append(D, max(np.abs(ecdf[x].y - norm_rv.cdf(ecdf[x].x))))
```



```
In [57]: x = np.linspace(1, N, N)
plt.plot(x, D)

plt.ylabel('$D_{n}\{ \}$')
plt.xlabel('$n$')
plt.title('$D_{n}\{ \}(n)$')
plt.show()
```

