

Домашнее задание 1 Вариант 5

Мосолков Евгений Николаевич БПИ196

Выборка:

24,24,25,10,59,20,14,22,14,84,21,22,86,35,21,20,15,24,29,12,30,37

Размер выборки: 22

А. Рассчитать 90% доверительный интервал для средней продолжительности, считая распределение признака нормальным.

Найдем выборочное среднее:

$$\bar{x} = \frac{1}{n} \sum_{j=0}^n x_j = \frac{648}{22} = 29.454546$$

Найдем дисперсию:

$$D_x = \frac{1}{n} \sum_{i=0}^n (x_i - \bar{x})^2 \approx 411.157025$$

Найдем среднеквадратичное отклонение:

$$\sigma_x = \sqrt{D_x} = \sqrt{411.157025} \approx 20.277007$$

Находим 90% доверительный интервал:

$$\bar{x} - t_i \frac{\sigma_x}{\sqrt{n}} < \alpha < \bar{x} + t_i \frac{\sigma_x}{\sqrt{n}}, \text{ где } t_i \text{ по таблицы Студента } a - 1.7207$$

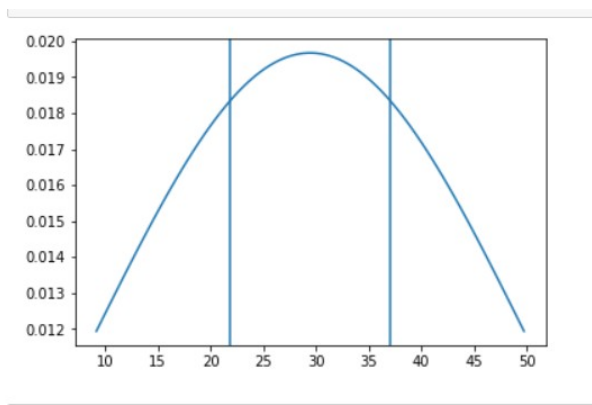
Считаем точность оценки:

$$\sigma = t_i \frac{\sigma_x}{\sqrt{n-1}} = 1.7207 \frac{20.277077}{\sqrt{22}} \approx 7.613763257$$

Считаем доверительный интервал. Дисперсия смешанная, значит под корнем не n, а n-1:

$$29.454546 - 7.613763257 < a < 29.454546 + 7.613763257$$

$$15,34482 < a < 22,65518$$

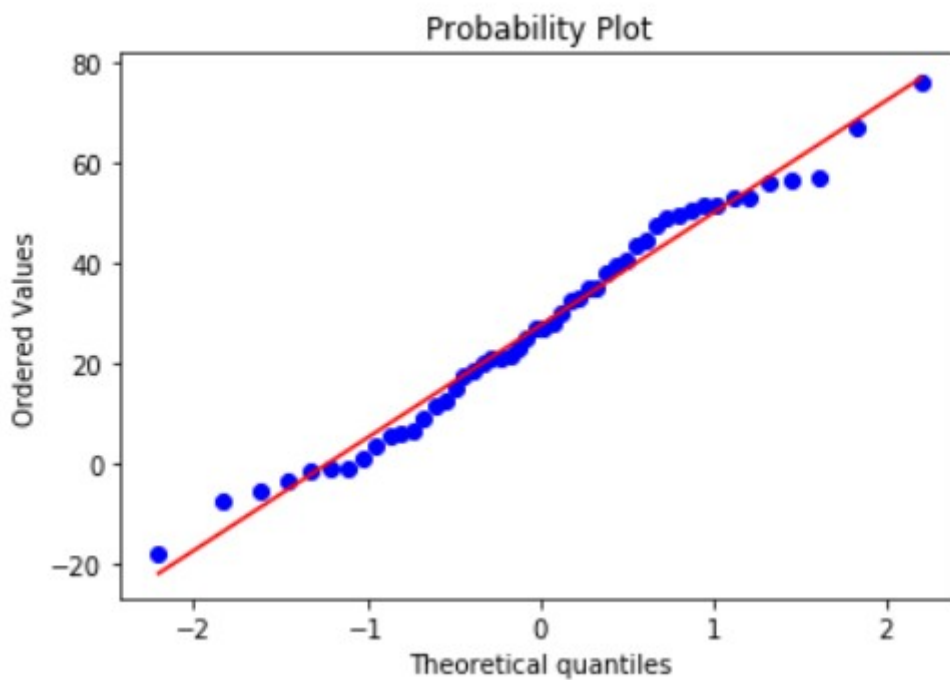


Пространство между вертикальными прямыми – это 90% доверительный интервал

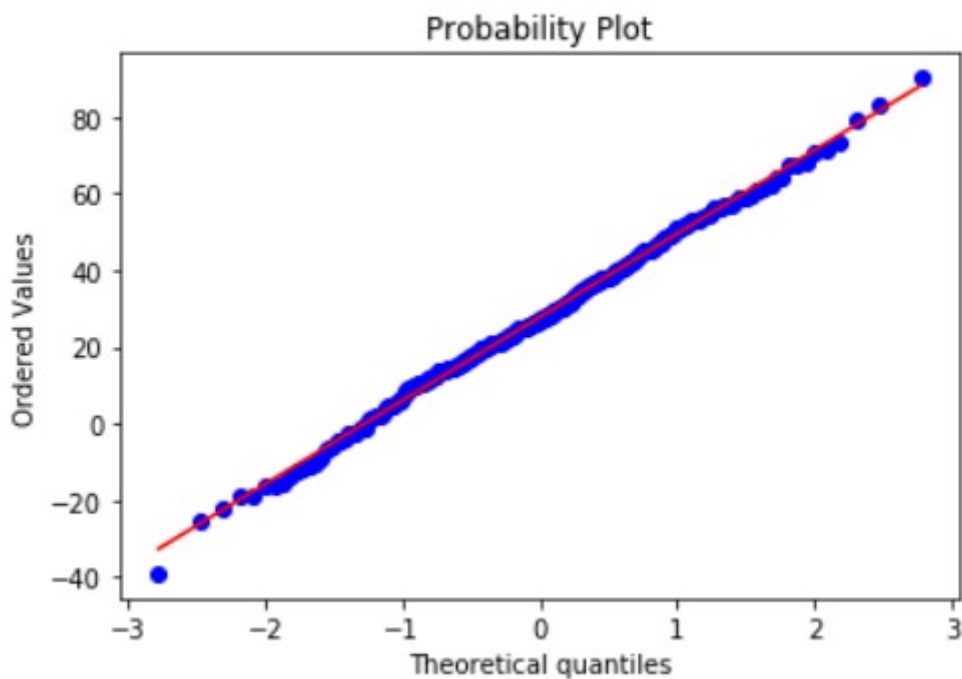
Б. Построить график квантиль-квантиль для нормального распределения с параметрами:

$$\bar{x}=29.454546$$

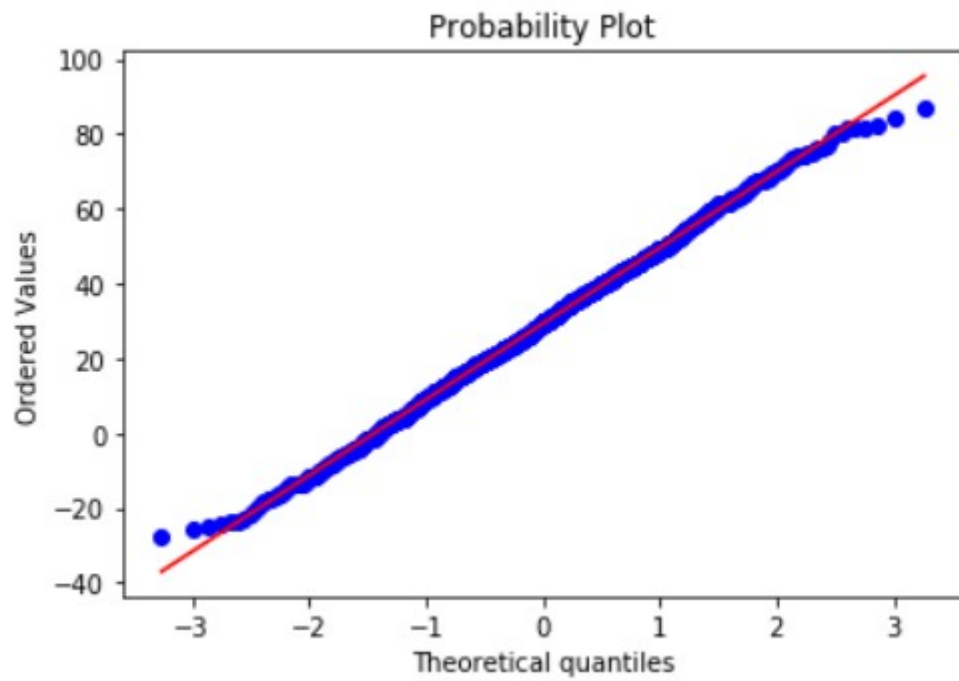
$$\sigma_x=20.277007$$



50 элементов – много значений лежит на прямой, значения лежащие не на прямой не сильно отдалены.



250 элементов – большинство точек лежат на прямой



1250 элементов – почти все точки лежат на прямой

Чем больше мы берем выборку, тем больше значений попадает на прямую, значит распределение соответствует нормальному закону.

В. Пользуясь методом bootstrap генерируем 1000 перевыборок исходной выборки

```
def bootstrap_points(data, count):
    i = np.random.randint(0, len(data), (count, len(data)))
    return data[i]
list_pts = bootstrap_points(data, 1000)
new_data = [x for i in list_pts for x in i]
print('Размер новой выборки', len(new_data))
print(new_data)
```

0, 24, 20, 33, 24, 10, 12, 22, 30, 37, 22, 21, 24, 21, 84, 20, 14, 23, 20, 20, 20, 13, 14, 23, 23, 80, 33, 21, 30, 24, 14, 2  
 0, 24, 21, 21, 29, 37, 14, 59, 12, 25, 29, 12, 86, 21, 37, 21, 20, 35, 14, 15, 35, 20, 37, 15, 22, 10, 24, 30, 20, 84, 35, 3  
 5, 22, 25, 59, 29, 21, 24, 35, 30, 21, 86, 21, 86, 14, 20, 29, 59, 84, 24, 86, 86, 24, 35, 59, 14, 59, 22, 24, 37, 14, 37, 2  
 4, 84, 20, 22, 15, 24, 37, 10, 30, 59, 37, 15, 37, 24, 35, 21, 86, 25, 37, 35, 21, 30, 86, 20, 20, 21, 29, 20, 24, 14, 14, 8  
 4, 29, 14, 59, 14, 22, 22, 24, 84, 21, 24, 10, 86, 20, 22, 24, 21, 12, 59, 35, 21, 10, 24, 12, 25, 20, 21, 21, 14, 12, 35, 1  
 0, 21, 20, 14, 22, 86, 84, 30, 84, 10, 25, 14, 24, 30, 84, 29, 12, 20, 35, 84, 24, 24, 14, 21, 15, 25, 20, 86, 86, 24, 21, 2  
 4, 29, 20, 24, 14, 21, 24, 12, 30, 35, 14, 14, 25, 25, 25, 59, 86, 29, 15, 12, 59, 86, 29, 29, 59, 84, 20, 14, 29, 22, 15, 1  
 4, 29, 84, 24, 29, 22, 59, 25, 10, 37, 25, 29, 35, 15, 21, 22, 24, 20, 30, 20, 14, 10, 20, 22, 15, 59, 20, 24, 14, 22, 21, 2  
 4, 15, 21, 22, 35, 86, 15, 10, 14, 37, 10, 35, 14, 35, 86, 20, 84, 14, 24, 86, 59, 15, 14, 29, 21, 25, 24, 22, 14, 20, 35, 2  
 2, 37, 15, 21, 24, 24, 25, 84, 10, 20, 14, 24, 21, 25, 84, 14, 24, 25, 14, 35, 21, 35, 30, 24, 22, 24, 37, 24, 10, 20, 22, 2  
 9, 35, 25, 24, 14, 20, 15, 37, 14, 25, 10, 30, 84, 22, 15, 29, 20, 22, 14, 10, 59, 24, 86, 30, 15, 10, 35, 30, 14, 20, 86, 1  
 4, 21, 15, 35, 35, 22, 14, 25, 35, 14, 86, 14, 84, 29, 15, 15, 12, 25, 22, 37, 22, 20, 20, 84, 22, 20, 22, 22, 86, 84, 20, 3  
 5, 35, 14, 35, 35, 22, 24, 20, 37, 35, 21, 25, 37, 35, 22, 20, 59, 25, 10, 25, 25, 12, 29, 14, 24, 24, 29, 20, 25, 10, 12, 2  
 2, 21, 25, 22, 30, 20, 35, 59, 15, 37, 22, 14, 10, 14, 24, 24, 24, 37, 21, 59, 24, 24, 21, 21, 15, 24, 14, 35, 86, 15, 37, 3  
 7, 35, 14, 12, 35, 14, 37, 24, 21, 84, 22, 37, 59, 24, 21, 21, 15, 24, 12, 86, 21, 30, 86, 84, 30, 10, 29, 84, 37, 22, 20, 2  
 4, 22, 24, 10, 86, 24, 86, 10, 20, 12, 24, 24, 86, 22, 30, 84, 86, 15, 24, 59, 86, 30, 84, 84, 24, 37, 20, 22, 25, 21, 29, 2  
 2, 24, 10, 14, 35, 12, 25, 35, 24, 24, 24, 35, 21, 25, 12, 37, 24, 24, 25, 25, 24, 24, 22, 21, 30, 30, 29, 22, 84, 35, 86, 2  
 0, 22, 84, 21, 24, 15, 86, 24, 14, 22, 29, 20, 12, 29, 25, 15, 14, 25, 22, 10, 21, 24, 22, 20, 22, 24, 12, 12, 22, 37, 21, 5  
 9, 37, 30, 59, 59, 14, 12, 25, 24, 21, 35, 14, 12, 20, 24, 59, 20, 35, 24, 14, 24, 59, 14, 24, 12, 15, 22, 15, 20, 21, 25, 2

Посчитаем выборочное среднее и среднеквадратическое отклонение новой выборки:

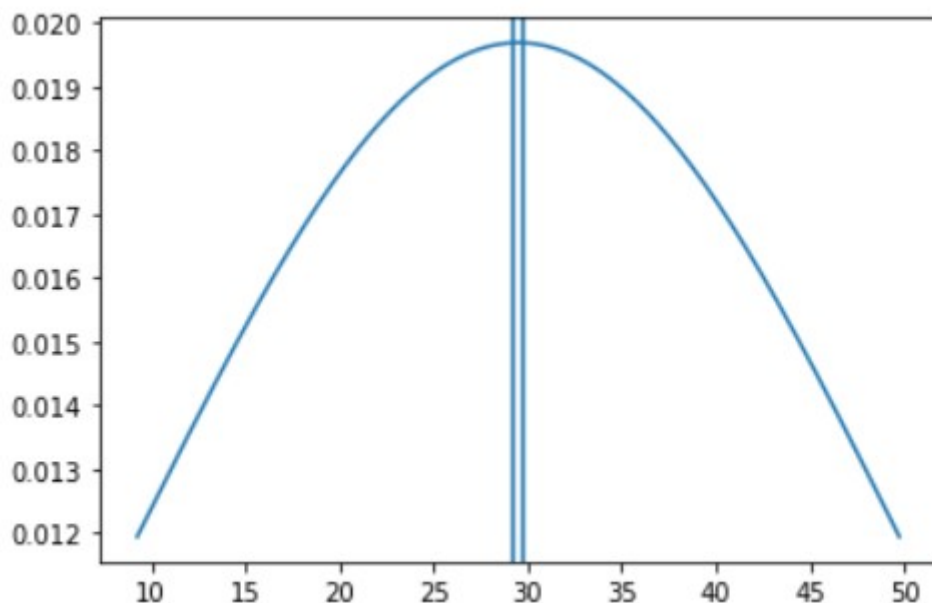
$$\bar{x} \approx 29.496727273$$

$$\sigma_x \approx 20.2610954612$$

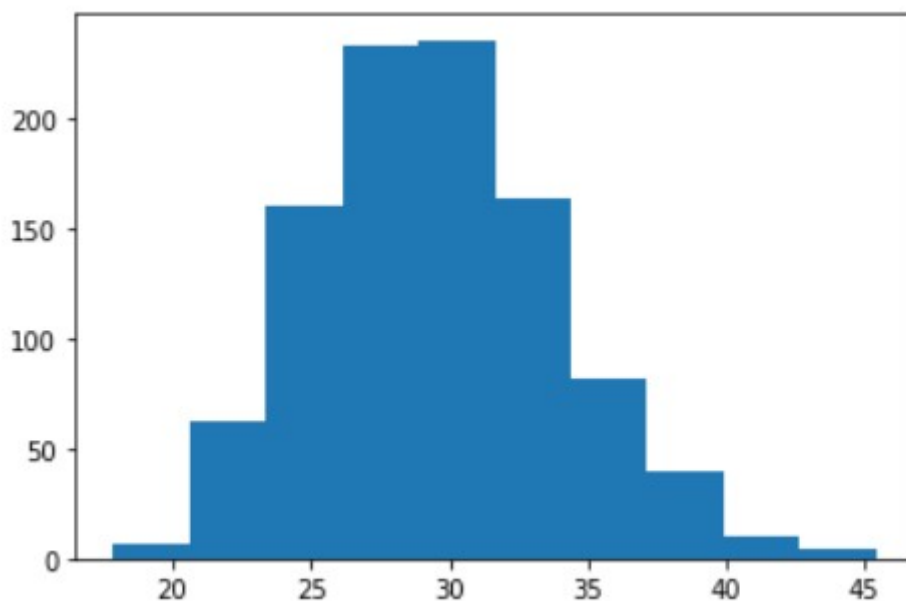
Строим доверительный интервал как в пункте А:

$$29.272025 < a < 29.72142$$

Доверительный интервал сократился. Очевидно что чем больше выборка – тем меньше доверительный интервал, при этом выборочное среднее и квадратическое отклонение изменились не сильно.



Г. Считаем среднее значение всех перевыборок и строим гистограмму на их основе



Распределение среднего в перевыборках является нормальным распределением. Наибольшая плотность - у выборочного среднего. Чем выше у значения разница с выборочным средним, тем меньше его плотность.