Отчёт по математической статистике.

По теме: «Доверительный интервал для нормального распределения».

Выполнила: Быкова К. А.

Студентка 433 группы

**Задача:** Смоделировать выборку объёма n=50 из нормального распределения N(a+1,d+2), где ab.cd – дата рождения. Т.е. в моём случае 04.05, следовательно, распределение будет N(1,7).

Построить интервалы при $\gamma = 0{,}95$ для:

А) Неизвестно а, $\delta^2 = 7$;

Б) Неизвестно а, если $\delta^2$ - неизвестно;

В) Неизвестно $\delta^2$, а $= 1$;

Г) Неизвестно $\delta^2$, если а – неизвестно.

Затем нарастить выборку до 100 элементов. Сделать вывод о том как, влияет объём выборки на длину интервала.

Для начала приведем определение доверительного интервала:

Будем говорить, что $(\theta_1, \theta_2) = 100 * \gamma\%$ - пара $\theta$. Если $\mathrm{P}(\theta_1 < \theta < \theta_2) = \gamma$.

Смоделируем выборку с помощью программы:

```
n=50;
x=normrnd(1,7,1,n)
sr_x=(1/n)*(sum(x))
S02=(1/n)*sum((x-sr_x).^2)
a=1;
Sa=(1/n)*sum((x-a).^2)
```

Выборка: x = -8.0004 -1.1981 2.7884 6.0486 -10.2917 9.5391 13.7505 -7.2412 -4.8050 -0.6105 -14.9490 8.2674 -6.5591 3.3356 7.9943 4.4388 6.3119 9.3749 4.0346 2.5261 -2.2457 -0.0868 -5.6387 -4.6998 -4.7039 2.2327 1.0639 -2.2474 -13.5613 -4.4499 -8.9879 6.5138 -8.5955 -6.1815 -0.9356 5.2187 0.1692 -1.9801 3.4729 -

3.5929 -7.9389 2.2678 -5.1424 5.0064 3.2388 7.7463 8.0935 16.3681 4.0013 0.2294

sr_x = 0.1878

S02 = 45.5565

Sa = 46.2162

А) В случае, когда неизвестно а, $\delta^2$ – известно, статистика имеет вид:

$$\sqrt{n} * \frac{\bar{x} - a}{\delta} \epsilon \mathrm{N}(0,1)$$

$$\mathrm{P}\left(C_1 < \sqrt{n} * \frac{\bar{x} - a}{\delta} < C_2\right) = 0.95$$

$$\mathrm{P}\left(\bar{x} - \frac{C_2 * \delta}{\sqrt{n}} < a < \bar{x} + \frac{C_2 * \delta}{\sqrt{n}}\right) = 0.95$$

По таблице нормального распределения находим $C_1, C_2$:

$$C_1 = -1.96, C_2 = 1.96$$

$$P\left(0.1878 - \frac{1.96 * 7}{\sqrt{50}} < a < 0.1878 + \frac{1.96 * 7}{\sqrt{50}}\right) = 0.95$$

Получаем:

$$P(-1.7525 < a < 2.1281) = 0.95$$

Т.е. получаем интервал а$\epsilon(-1.7525; 2.1281)$.

Б) В случае, когда неизвестно а, $\delta^2$ – неизвестно, статистика имеет вид:

$$\sqrt{n} * \frac{\bar{x} - a}{S_0} \epsilon \tau_{n-1}$$

$$\mathrm{P}\left(\bar{x} - \frac{C_2 * \mathrm{S}_0}{\sqrt{n}} < a < \bar{x} + \frac{C_2 * \mathrm{S}_0}{\sqrt{n}}\right) = 0.95$$

$$S_0 = \sqrt{S02} = \sqrt{45.5565} = 6.7496$$

Т.к. степень свободы нашего распределения довольно большая, мы можем воспользоваться тем, что при $n \to \infty$ распределение Стьюдента стремиться к нормальному распределению. Следовательно, возьмём $C_1 = -1.96, C_2 =$

$= 1.96$.

$$P\left(0.1878 - \frac{1.96 * 6.7496}{\sqrt{50}} < a < 0.1878 + \frac{1.96 * 6.7496}{\sqrt{50}}\right) = 0.95$$

Получаем:

$$P(-1.6831 < a < 2.0587) = 0.95$$

Т.е. получаем интервал a$\epsilon(-1.6831; 2.0587)$.

В) В случае, когда неизвестно $\delta^2$, a-известно $a = 1$, статистика имеет вид:

$$\frac{n * S_0^2(a)}{\delta^2} \epsilon \chi_n^2$$

$$P\left(\frac{n * S_0^2(a)}{C_2} < \delta^2 < \frac{n * S_0^2(a)}{C_1}\right) = 0.95$$

Найдём значения $C_1, C_2$ по таблице распределения $\chi_n^2$:

$$C_1 = 32.3574, C_2 = 71.4202$$

$$P\left(\frac{50 * 46.2162}{32.3574} < \delta^2 < \frac{50 * 46.2162}{71.4202}\right) = 0.95$$

$$P(32.3552 < \delta^2 < 71.4152) = 0.95$$

Т.е. получаем интервал $\delta^2 \epsilon(32.3552; 71.4152)$.

Г) В случае, когда неизвестно $\delta^2$, и a – неизвестно, статистика имеет вид:

$$\frac{(n-1) * S_0^2(a)}{\delta^2} \epsilon \chi_{n-1}^2$$

$$P\left(\frac{(n-1) * S_0^2(a)}{C_2} < \delta^2 < \frac{(n-1) * S_0^2(a)}{C_1}\right) = 0.95$$

Найдём значения $C_1, C_2$ по таблице распределения $\chi_{n-1}^2$:

$$C_1 = 31.5549, C_2 = 70.2224$$

$$P\left(\frac{49 * 46.2162}{70.2224} < \delta^2 < \frac{49 * 46.2162}{31.5549}\right) = 0.95$$

$$P(32.907 < \delta^2 < 73.6978) = 0.95$$

Т.е. получаем интервал $\delta^2 \epsilon(32.907; 73.6978)$.

Наращиваем выборку на 50 элементов:

```
n=100;
x=normrnd(1,7,1,n);
sr_x=(1/n)*(sum(x))
```

```
S02=(1/n)*sum((x-sr_x).^2)
a=1;
Sa=(1/n)*sum((x-a).^2)
sr_x =1.3057
S02 = 53.7068
Sa =53.8003
```

А) P $\left(\bar{x} - \frac{C_2 * \delta}{\sqrt{n}} < a < \bar{x} + \frac{C_2 * \delta}{\sqrt{n}}\right) = 0.95$

По таблице нормального распределения находим $C_1, C_2$:

$$C_1 = -1.96, C_2 = 1.96$$

$$P\left(1.3057 - \frac{1.96 * 7}{\sqrt{50}} < a < 1.3057 + \frac{1.96 * 7}{\sqrt{50}}\right) = 0.95$$

Получаем:

$$P(-0.6346 < a < 3.246) = 0.95$$

Т.е. получаем интервал а$\epsilon(-0.6346; 3.246)$.

Б) P $\left(\bar{x} - \frac{C_2 * \mathrm{S}_0}{\sqrt{n}} < a < \bar{x} + \frac{C_2 * \mathrm{S}_0}{\sqrt{n}}\right) = 0.95$

$$S_0 = \sqrt{S02} = \sqrt{53.7068} = 7.3285$$

Т.к. степень свободы нашего распределения довольно большая, мы можем воспользоваться тем, что при $n \to \infty$ распределение Стьюдента стремиться к нормальному распределению. Следовательно, возьмём $C_1 = -1.96, C_2 =$

$= 1.96$.

$$P\left(1.3057 - \frac{1.96 * 7.3285}{\sqrt{100}} < a < 1.3057 + \frac{1.96 * 7.3285}{\sqrt{100}}\right) = 0.95$$

Получаем:

$$P(-0.1307 < a < 2.7421) = 0.95$$

Т.е. получаем интервал а$\epsilon(-0.1307; 2.7421)$.

В)P $\left(\frac{\mathrm{n} * \mathrm{S}_0^2(a)}{C_2} < \delta^2 < \frac{\mathrm{n} * \mathrm{S}_0^2(a)}{C_1}\right) = 0.95$

Найдём значения $C_1, C_2$ по таблице распределения $\chi_n^2$:

$$\mathrm{C}_1 = 77.92, C_2 = 129.561$$

$$\mathrm{P}\left(\frac{100 * 53.8003}{129.561} < \delta^2 < \frac{100 * 53.8003}{77.929}\right) = 0.95$$

$$\mathrm{P}(41.525 < \delta^2 < 69.037) = 0.95$$

Т.е. получаем интервал $\delta^2 \epsilon (41.525; 69.037)$.

Г)$P\left(\frac{(n-1)*S_0^2(a)}{C_2} < \delta^2 < \frac{(n-1)*S_0^2(a)}{C_1}\right) = 0.95$

Найдём значения $C_1, C_2$ по таблице распределения $\chi_{n-1}^2$:

$$C_1 = 77.046, C_2 = 128.422$$

$$P\left(\frac{99 * 53.8003}{128.422} < \delta^2 < \frac{99 * 53.8003}{77.046}\right) = 0.95$$

$$P(41.474 < \delta^2 < 69.13) = 0.95$$

Т.е. получаем интервал $\delta^2 \epsilon (41.474; 69.13)$.

**Вывод:**

А)

В первом случае а$\epsilon(-1.7525; 2.1281)$ интервал =3.8806

Во втором случае а$\epsilon(-0.6346; 3.246)$ интервал =3.8806

Б)

В первом случае а$\epsilon(-1.6831; 2.0587)$ интервал =3.7418

Во втором случае а$\epsilon(-0.1307; 2.7421)$ интервал =2.8728

В)

В первом случае $\delta^2 \epsilon (32.3552; 71.4152)$ интервал =39.06

Во втором случае $\delta^2 \epsilon (41.525; 69.037)$ интервал =27.512

Г)

В первом случае $\delta^2 \epsilon (32.907; 73.6978)$ интервал =40.7908

Во втором случае $\delta^2 \epsilon (41.474; 69.13)$ интервал =27.656

Следовательно, с увеличением объёма выборки доверительный интервал уменьшается.