Lecture 14
Interval estimators of parameters

Pointwise estimators of an unknown parameter are good as initial results of processing observations. Their disadvantage is that it is unknown with which precision they give an estimated value. For samples of small size the question on precision of estimators is very significant, since it can be a great discrepancy between θ and θ* in this case. In addition, at solving practical tasks it is often required to determine the reliability of these estimators. Therefore the problem on approximation of the parameter θ by a whole interval (θ_1^*, θ_2^*) (not only by one number) arises.

An estimator of an unknown parameter is *interval* if it is determined by two numbers – endpoints of an interval.

The task of interval estimation can be formulated as follows: construct a numeric interval (θ_1^*, θ_2^*) on sample data on which we can say that the exact value of estimated parameter is within this interval with probability chosen beforehand.

An interval is said to be *confident* (or *confidence interval)* if it covers a given parameter with a given *reliability* (*confidence probability*) γ.

The quantity γ is chosen beforehand, its choice depends on specifically solved task. For example, the confidence level of an air passenger to the reliability of an aircraft, obviously, should be greater than the buyer's confidence level to the reliability of a TV set, lamp, toy and etc. The reliability γ is chosen equal to 0,9; 0,95; 0,99 or 0,999. Then it is practically reliably finding the parameter θ in the confidence interval.

1. An interval estimator (with reliability γ) of mathematical expectation *a* of a normally distributed quantitative attribute *X* on sample mean $\bar{x}_s$ at *known* mean square deviation σ of a parent population is the confidence interval

$$\bar{x}_s - t\cdot(\sigma/\sqrt{n}) < a < \bar{x}_s + t\cdot(\sigma/\sqrt{n})$$

where $t\cdot(\sigma/\sqrt{n}) = \delta$ is the accuracy of estimation, *n* is sample size, *t* is value of argument of Laplace function *Φ(t)* (Appendix 1) for which *Φ(t)* = γ/2.

If σ is *unknown* (and sample size *n* < 30) then

$$\bar{x}_s - t_\gamma\cdot(s/\sqrt{n}) < a < \bar{x}_s + t_\gamma\cdot(s/\sqrt{n})$$

where *s* is revised sample mean square deviation, $s^2 = \frac{n}{n-1}\cdot D_s$; t_γ can be found by Table (Appendix 2) on given *n* and γ.

2. An interval estimator (with reliability γ) of mean square deviation σ of a normally distributed quantitative attribute *X* on revised sample mean square deviation *s* is the confidence interval

$$s\cdot(1-q) < \sigma < s\cdot(1+q) \quad (for \quad q<1),$$

$$0 < \sigma < s\cdot(1+q) \quad (for \quad q>1),$$

where *q* is found by Table (Appendix 3) on given *n* and γ.

3. An interval estimator (with reliability γ) of unknown probability *p* of a binomial distribution on relative frequency *w* is the confidence interval (with approximated endpoints p_1 and p_2)

$$p_1 < p < p_2,$$

where

$$p_1 = \frac{n}{t^2+n}\cdot\left[w+\frac{t^2}{2n}-t\sqrt{\frac{w(1-w)}{n}+\left(\frac{t}{2n}\right)^2}\right],$$

$$p_2 = \frac{n}{t^2+n}\cdot\left[w+\frac{t^2}{2n}+t\sqrt{\frac{w(1-w)}{n}+\left(\frac{t}{2n}\right)^2}\right],$$

where n is the general number of trials; m is the number of appearances of an event; w is the relative frequency that is equal to ratio m/n; t is the value of argument of Laplace function (Appendix 1) for which $\Phi(t) = \gamma/2$ (γ is a given reliability).

Remark. For large values n (order of hundreds) we can take as approximated boundaries of the confidence interval

$$p_1 = w - t\sqrt{\frac{w(1-w)}{n}}, \quad p_2 = w + t\sqrt{\frac{w(1-w)}{n}}\ .$$

Example. Find the confidence interval for an estimator with reliability 0,95 of unknown mathematical expectation a of a normally distributed attribute X of a parent population if the parent mean square deviation $\sigma = 5$, sample mean $\bar{x}_s = 14$ and the sample size $n = 25$.

Solution: It is required to find the confidence interval

$$\bar{x}_s - t\cdot\frac{\sigma}{\sqrt{n}} < a < \bar{x}_s + t\cdot\frac{\sigma}{\sqrt{n}} \qquad (*)$$

All the quantities but t are known. Find t from the following formula: $\Phi(t) = 0{,}95/2 = 0{,}475$.
By Table (Appendix 1) we find $t = 1{,}96$.
Replacing $t = 1{,}96$, $\bar{x}_s = 14$, $\sigma = 5$, $n = 25$ in (*) we finally obtain the required confidence interval: $12{,}04 < a < 15{,}96$.

Example. Find minimal size of a sample at which with reliability 0,975 the accuracy of an estimator of mathematical expectation a of a parent population on sample mean is equal to $\delta = 0{,}3$ if it is known that the mean square deviation of a normally distributed parent population $\sigma = 1{,}2$.

Solution: We will use the formula determining the accuracy of estimator of mathematical expectation of a parent population on sample mean: $\delta = t\cdot(\sigma/\sqrt{n})$.

We have the following: $n = t^2\sigma^2/\delta^2$ (*)

By the hypothesis, $\gamma = 0{,}975$; consequently, $\Phi(t) = 0{,}975/2 = 0{,}4875$. By Table (Appendix 1) we find $t = 2{,}24$. Replacing $t = 2{,}24$, $\sigma = 1{,}2$ and $\delta = 0{,}3$ in (*), we obtain the required size of a sample: $n = 81$.

Example. A sample of size $n = 10$ has been extracted from a parent population:

variant	x_i	–2	1	2	3	4	5
frequency	n_i	2	1	2	2	2	1

Estimate with reliability 0,95 the mathematical expectation a of a normally distributed attribute of a parent population on sample mean by the confidence interval.

Solution: Find the sample mean and revised mean square deviation by the formulas:

$$\bar{x}_s = \frac{\sum n_i x_i}{n}, \quad s = \sqrt{\frac{\sum n_i (x_i - \bar{x}_s)^2}{n-1}}$$

Replacing the data of the example in these formulas, we obtain: $\bar{x}_s = 2$, $s = 2{,}4$.

Find t_γ. By using Table (Appendix 2) on γ = 0,95 and n = 10 we find t_γ= 2,26.
Find the required confidence interval:

$$\bar{x}_s - t_\gamma \cdot (s/\sqrt{n}) < a < \bar{x}_s + t_\gamma \cdot (s/\sqrt{n})$$

Replacing $\bar{x}_s = 2,\ \ s = 2{,}4,\ t_\gamma = 2{,}26,\ n = 10,$ we obtain the required confidence interval

$$0{,}3 < a < 3{,}7$$

covering the unknown mathematical expectation a with reliability 0,95.

<u>*Example.*</u> By data of a sample of size n = 16 from a parent population the following has been found: the revised mean square deviation s = 1 of a normally distributed quantitative attribute. Find the confidence interval covering parent mean square deviation σ with reliability 0,95.
Solution: The task is reduced to finding the confidence interval

$$s \cdot (1-q) < \sigma < s \cdot (1+q) \quad (for \quad q < 1),$$

or (*)

$$0 < \sigma < s \cdot (1+q) \quad (for \quad q > 1).$$

Using the data γ = 0,95 and n = 16 by Table (Appendix 3) we find q = 0,44. Since q < 1, replacing s = 1, q = 0,44 in (*), we obtain the required confidence interval

$$0{,}56 < \sigma < 1{,}44.$$

<u>*Example.*</u> 12 measurements of some physical quantity have been made by one device (without systematic mistakes), and the revised mean square deviation s of random mistakes appeared equal to 0,6. Find the accuracy of the device with reliability 0,99. It is supposed that results of measurements are normally distributed.
Solution: The accuracy of the device is characterized by mean square deviation of random mistakes of measuring. Therefore the task is reduced to finding the confidence interval covering σ with the given reliability γ = 0,99:

$$s \cdot (1-q) < \sigma < s \cdot (1+q) \quad (for \quad q < 1), \qquad (*)$$

Using the data γ = 0,99 and n = 12 by Table (Appendix 3) we find q = 0,9. Since q < 1, replacing s = 0,6; q = 0,9 in (*), we obtain the required confidence interval

$$0{,}06 < \sigma < 1{,}14.$$

<u>*Example.*</u> Independent trials are made with identical but unknown probability p of appearance of an event A in each trial. Find the confidence interval for estimation of probability p with reliability 0,95 if the event A has appeared 15 times in 60 trials.
Solution: By the hypothesis n = 60, m = 15, γ = 0,95. Find the relative frequency of appearance of the event A*:* $w = m/n$ = 15/60 = 0,25.
Find t from the formula $\Phi(t) = \gamma/2$ = 0,95/2 = 0,475. By Table of Laplace function (Appendix 1) we find t = 1,96.
Find boundaries of the required confidence interval:

$$p_1 = \frac{n}{t^2+n} \cdot \left[w + \frac{t^2}{2n} - t\sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2} \right],$$

$$p_2 = \frac{n}{t^2+n} \cdot \left[w + \frac{t^2}{2n} + t\sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2} \right],$$

Replacing n = 60, w = 0,25; t = 1,96 in these formulas, we obtain p_1 = 0,16; p_2 = 0,37.
Thus, the required confidence interval 0,16 < p < 0, 37.

Example. An experimental slot machine providing an appearance of a winning only once of 100 throwings a coin in the machine has been made. 400 trials have been made to check suitability of the slot machine, and the winning has appeared 5 times in these trials. Find the confidence interval covering unknown probability of appearance of the winning with reliability $\gamma = 0,999$ (slot machine – игровой автомат; winning – выигрыш; suitability – пригодность).
Solution: Find the relative frequency of appearance of the winning: $w = m/n = 5/400 = 0,0125$. Find t from the formula $\Phi(t) = \gamma/2 = 0,999/2 = 0,4995$. By Table of Laplace function (Appendix 1) we find $t = 3,3$.
Taking in account that $n = 400$ is great, we use the approximate formulas for finding boundaries of the confidence interval: $p_1 = w - t\sqrt{\dfrac{w(1-w)}{n}}, \quad p_2 = w + t\sqrt{\dfrac{w(1-w)}{n}}$.
Replacing $w = 0,0125$; $t = 3,3$; $n = 400$ in these formulas, we obtain $p_1 = -0,0058$, $p_2 = 0,0308$. Thus, the required confidence interval $0 < p < 0,0308$.

Glossary

discrepancy – расхождение; **reliability** – надежность
confidence interval – доверительный интервал
accuracy – точность; **slot machine** – игровой автомат

Exercises for Seminar 14

14.1. Find the confidence interval for an estimator with reliability 0,99 of unknown mathematical expectation a of a normally distributed attribute X of a parent population if the parent mean square deviation σ, sample mean $\bar{x}_s$ and sample size n are known: $\sigma = 4$, $\bar{x}_s = 10,2$; $n = 16$.
The answer: $7,63 < a < 12,77$.

14.2. Five equally exact measurements of distance between a gun and a target have been made by the same device with the mean square deviation of random mistakes of measurements $\sigma = 40$ *m*. Find the confidence interval for an estimator of true distance a to the target with reliability $\gamma = 0,95$, knowing that the arithmetic mean of results of measurements $\bar{x}_s = 2000$ *m*. It is supposed that results of measurements are normally distributed.
The answer: $1964,94 < a < 2035,06$.

14.3. An automatic machine stamps cylinders. The sample mean of diameters of made cylinders has been calculated on a sample $n = 100$. Find the accuracy δ with reliability 0,95 on which the sample mean estimates the mathematical expectation of diameters of making cylinders, knowing that their mean square deviation $\sigma = 2$ *mm*. It is supposed that diameters of cylinders are normally distributed.
The answer: $\delta = 0,392$ *mm*.

14.4. Find minimal size of a sample at which with reliability 0,925 the accuracy of estimator of mathematical expectation of a normally distributed parent population on sample mean is equal to 0,2 if the mean square deviation of the parent population $\sigma = 1,5$.
The answer: $n = 179$.

14.5. A sample of size $n = 12$ has been extracted from a parent population:

variant	x_i	–0,5	–0,4	–0,2	0	0,2	0,6	0,8	1	1,2	1,5
frequency	n_i	1	2	1	1	1	1	1	1	2	1

Estimate with reliability 0,95 the mathematical expectation a of a normally distributed attribute of a parent population on sample mean by the confidence interval.
The answer: $-0,04 < a < 0,88$.

14.6. By data of nine independent equally exact measurements of some physical quantity the following has been found: the arithmetic mean of results of measurements $\bar{x}_s = 30{,}1$ and revised mean square deviation *s* = 6. Estimate true value of the measured quantity by the confidence interval with reliability γ = 0,99. It is supposed that results of measurements are normally distributed.

14.7. By data of a sample of size *n* from a parent population of a normally distributed quantitative attribute the following has been found: the revised mean square deviation *s*. Find the confidence interval covering the parent mean square deviation σ with reliability 0,999 if *n* = 10, *s* = 5,1.
The answer: $0 < \sigma < 14{,}28$.

14.8. 10 measurements of some physical quantity have been made by one device (without systematic mistakes), and the revised mean square deviation *s* of random mistakes appeared equal to 0,8. Find the accuracy of the device with reliability 0,95. It is supposed that results of measurements are normally distributed.
The answer: $0{,}28 < \sigma < 1{,}32$.

14.9. Independent trials are made with identical but unknown probability *p* of appearance of an event *A* in each trial. Find the confidence interval for estimation of probability *p* with reliability 0,99 if the event *A* has appeared 60 times in 100 trials.
The answer: $0{,}47 < p < 0{,}71$.

14.10. An event *A* has appeared 270 times in 360 trials in each of which the probability of appearance of the event is constant and unknown. Find the confidence interval covering the unknown probability *p* with reliability 0,95.
The answer: $0{,}705 < p < 0{,}795$.

14.11. 100 refusals have been registered when 1000 elements were testing. Find the confidence interval covering an unknown probability *p* of refusal of an element with reliability 0,95.

Exercises for Homework 14

14.12. Find the confidence interval for an estimator with reliability 0,99 of unknown mathematical expectation *a* of a normally distributed attribute *X* of a parent population if the parent mean square deviation σ, sample mean $\bar{x}_s$ and sample size *n* are known: $\sigma = 5$, $\bar{x}_s = 16{,}8$; *n* = 25.
The answer: $14{,}23 < a < 19{,}37$.

14.13. A sample of big butch of electro lamps contains 100 lamps. The average durability of burning a lamp of the sample is 1000 hours. Find the confidence interval with reliability 0,95 for average durability *a* of burning a lamp of all the butch if it is known that the mean square deviation of durability of burning a lamp σ = 40 h. it is supposed that the durability of burning lamps is normally distributed.
The answer: $992{,}16 < a < 1007{,}84$.

14.14. By data of 16 independent equally exact measurements of some physical quantity the following has been found: the arithmetic mean of results of measurements $\bar{x}_s = 42{,}8$ and revised mean square deviation *s* = 8. Estimate true value of the measured quantity with reliability γ = 0,999.
The answer: $34{,}66 < a < 50{,}94$.

14.15. By data of a sample of size n from a parent population of a normally distributed quantitative attribute the following has been found: the revised mean square deviation s. Find the confidence interval covering the parent mean square deviation σ with reliability 0,999 if $n = 50$, $s = 14$.
The answer: $7{,}98 < \sigma < 20{,}02$.

14.16. 300 trials have been made in each of which an unknown probability p of appearance of an event A is constant. The event A has appeared 250 times in these trials. Find the confidence interval covering the unknown probability p with reliability 0,95.
The answer: $0{,}78 < p < 0{,}87$.

14.17. 32 non-standard details have appeared in a batch of 250 details made by an automatic machine. Find the confidence interval covering with reliability 0,99 an unknown probability p of making a non-standard detail by the machine.
The answer: $0{,}07 < p < 0{,}18$.

14.18. 100 refusals have been registered when 1000 elements were testing. Find the confidence interval covering an unknown probability p of refusal of an element with reliability 0,99.
The answer: $0{,}076 < p < 0{,}124$.