4.4 Интевали на доверба

Задача 4.41. При 100 независни гаѓања стрелецот ја погодил целта 50 пати. Најди 99% интервал на доверба за непознатата веројатност за погодок на целта при едно гаѓање во метата.

Решение. Обележјето $X = I_A$, каде настанот A е целта е погодена при едно гаѓање на метата, има $\mathcal{B}(1,p)$ распределба, каде $0 < p < 1$ е непознат параметар. Нека $(X_1, ..., X_n)$ е примерок кој одговара на обележјето X, тогаш интервалот на доверба за p со веројатност на доверба $1-\alpha$ е

$$I_p = \Big(\frac{\mu}{n} - u_{(1-\alpha)/2}\sqrt{\frac{\mu}{n^2}\Big(1-\frac{\mu}{n}\Big)}, \frac{\mu}{n} + u_{(1-\alpha)/2}\sqrt{\frac{\mu}{n^2}\Big(1-\frac{\mu}{n}\Big)}\Big),$$

каде $\mu = \sum_{i=1}^n X_i$. За дадениот примерок важи $n = 100$, $\mu = 50$ и $1-\alpha = 0,99$, од каде $u_{(1-\alpha)/2} = u_{0,495}$ е такво да $\Phi_0(u_{0,495}) = 0,495$, па од таблица имаме $u_{0,495} = 2,575$. Тогаш, бараниот 99% интевал на доверба за p е $I_p = (0,37125; 0,62875)$.

Задача 4.42. Обележјето X на една популација има $\mathcal{U}(0, 1+\theta)$ распределба, каде $\theta > 0$ е непознат параметар. Во примерок со големина $n = 200$ кој одговара на обележејот X има 150 елементи кои се помали од 1. Врз основа на тој примерок, најди 95% интервал на доверба за веројатнста $P\{X < 1\}$, а потоа и 95% интервал на доверба за θ.

Решение. Непознатата веројатност ја означуваме со $p = P\{X < 1\}$. Тогаш, $Y = I_{\{X<1\}}$ има $\mathcal{B}(1,p)$ распределба. Нека $(X_1, ..., X_n)$ е примерок кој одговара на обележјето X, тогаш $(Y_1, ..., Y_n)$, каде $Y_i = I_{\{X_i<1\}}$ е примерок кој одговара на обележјето Y. Дадено е дека $n = 200$ и дека 150 елементи од примерокот $(X_1, ..., X_n)$ се помали од 1, што значи дека 150 елементи од примерокот $(Y_1, ..., Y_n)$ се 1-ци, а останатите се 0-ли т.е. $\mu = \sum_{i=1}^{200} Y_i = 150$. Тогаш, интервалот на доверба за p со веројатност на доверба $1-\alpha = 0,95$ е

$$I_p = \Big(\frac{\mu}{n} - u_{(1-\alpha)/2}\sqrt{\frac{\mu}{n^2}\Big(1-\frac{\mu}{n}\Big)}, \frac{\mu}{n} + u_{(1-\alpha)/2}\sqrt{\frac{\mu}{n^2}\Big(1-\frac{\mu}{n}\Big)}\Big),$$

каде $\mu = \sum_{i=1}^n Y_i$. За $n = 200$, $\mu = 150$ и $u_{(1-\alpha)/2} = u_{0,475}$ за кое $\Phi_0(u_{0,475}) = 0,475$, па затоа $u_{0,475} = 1,96$, од каде имаме дека 95% интервал на доверба за p е $I_p = (0,689988; 0,810012)$.

Сега, од

$$p = P\{X < 1\} = \int_{-\infty}^1 p_X(x,\theta)\,dx = \int_0^1 \frac{1}{1+\theta}\,dx = \frac{1}{1+\theta},$$

и од тоа што $I_p=(0,689988;0,810012)$ е 95% интервал на доверба за p, имаме

$$\begin{aligned}0,95 &= P\{0,689988<p<0,810012\}=P\{0,689988<\frac{1}{1+\theta}<0,810012\}=\\ &= P\{1,23455<1+\theta<1,449301\}=P\{0,23455<\theta<0,449301\},\end{aligned}$$

односно $I_\theta=(0,23455;0,449301)$ е бараниот 95% интервал на доверба за θ.

Задача 4.43. Измерени се висините на 10 студенти на математика и добиени се следните вредности во cm: 165, 182, 173, 154, 157, 162, 168, 176, 168, 170. Под претпоставка дека X-висината на еден студент на математика има нормална распределба, најди 90% интервал на доверба за очекуваната висина на студентите на математика.

Решение. Обележјето X-висината на еден студент на математика има $\mathcal{N}(m,\sigma^2)$ распределба, при што m и σ^2 се непознати параметри. Нека $(X_1,...,X_n)$ е примерок кој одговара на обележјето X. Се бара интервал на доверба за m, кога σ^2 не е позната. Во тој случај имаме дека, интервалот на доверба за m со веројатност на доверба $1-\alpha$ е

$$I_m=\Big(\overline{X}_n-t_{n-1,\alpha}\,\frac{\overline{S}_n}{\sqrt{n-1}},\overline{X}_n+t_{n-1,\alpha}\,\frac{\overline{S}_n}{\sqrt{n-1}}\Big),$$

каде $\overline{X}_n=\frac{1}{n}\sum_{i=1}^n X_i$, $\overline{S}_n^2=\frac{1}{n}\sum_{i=1}^n X_i^2-\overline{X}_n^2$ и $t_{n-1,\alpha}$ е таков да $P\{|T_{n-1}|>t_{n-1,\alpha}\}=\alpha$, каде T_{n-1} е случјана променлива со студентова распределба со $(n-1)$ степени на слобода. За дадените податоци имаме $n=10$, потоа

$$\overline{X}_n=\frac{1}{10}(165+182+173+154+157+162+168+176+168+170)=167,5,$$

$$\overline{S}_n^2=\frac{1}{10}(165^2+182^2+173^2+154^2+157^2+162^2+168^2+176^2+168^2+170^2)-167,5^2=64,85,$$

од каде $\overline{S}_n=8,05295$. Веројатноста на доверба на бараниот интервал е $1-\alpha=0,9$, од каде $\alpha=0,1$, па од таблицата за студентова распределба наоѓаме $t_{n-1,\alpha}=t_{9;0,1}=1,833$. Така, добиваме дека врз основа на дадените податоци, 90% интервал на доверба за очекуваната висина m е $I_m=(162,58;172,42)$.

Задача 4.44. Обележјето X на една популација има $\mathcal{N}(m,0,1^2)$ распределба. Одреди ја најмалата големина n на примерокот $(X_1,...,X_n)$ за да должината на 95% интервал на доверба за m да не е поголема од 0,05.

Решение. Обележјето X има $\mathcal{N}(m,0,1^2)$ распределба, каде m е непознат параметар. Нека $(X_1,...,X_n)$ е примерок кој одговара на обележјето X. Интервал на доверба за m кога $\sigma^2=0,1^2$ е позната, со веројатност на доверба $1-\alpha$ е

$$I_m=\Big(\overline{X}_n-u_{(1-\alpha)/2}\,\frac{\sigma}{\sqrt{n}},\overline{X}_n+u_{(1-\alpha)/2}\,\frac{\sigma}{\sqrt{n}}\Big),$$

каде $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ и $u_{(1-\alpha)/2}$ е таков да $\Phi_0(u_{(1-\alpha)/2}) = (1-\alpha)/2$. Веројатноста на доверба е $1-\alpha = 0,95$, од каде $u_{(1-\alpha)/2} = u_{0,475}$ е таков да $\Phi_0(u_{0,475}) = 0,475$, т.е. $u_{0,475} = 1,96$. Тогаш, за должината на интервалот имаме

$$dI_m = 2 \cdot u_{(1-\alpha)/2}\,\frac{\sigma}{\sqrt{n}} = 2 \cdot 1,96 \cdot \frac{0,1}{\sqrt{n}} = \frac{0,392}{\sqrt{n}}.$$

Се бара најмалата вредност за n за која $dI_m \leq 0,05$, односно $\frac{0,392}{\sqrt{n}} \leq 0,05$. Решение на последната неравенка е $n \geq 61,4656$ и бидејќи n е природен број, следи дека најмалата вредност на примерокот е $n = 62$.

Задача 4.45. За време на еден студенцки натпревар во трчање на 100 метри, случајно избрани 30 натпреварувачи ги постигнале следните резултати во секунди:

$$15, 12, 14, 15, 18, 16, 12, 17, 15, 23, 25, 11, 16, 14, 10$$
$$17, 16, 13, 16, 19, 15, 22, 12, 18, 15, 20, 14, 19, 16, 13.$$

Под претпоставка дека времето во секунди за кое еден студент - натпреварувач истрчува 100 метри, има нормална $\mathcal{N}(16, \sigma^2)$ распределба, врз основа на дадените податоци, најди 90% интервал на доверба за дисперзијата σ^2.

Решение. Обележјето X има $\mathcal{N}(16, \sigma^2)$ распределба, каде σ^2 е непознат параметар. Нека $(X_1, ..., X_n)$ е примерок кој одговара на обележјето X. Тогаш, интервал на доверба за дисперзијата σ^2 со веројатност на доверба $1-\alpha$ е

$$I_{\sigma^2} = \Big(\frac{\sum_{i=1}^{n}(X_i - m)^2}{\chi^2_{n,\alpha/2}}, \frac{\sum_{i=1}^{n}(X_i - m)^2}{\chi^2_{n,1-\alpha/2}}\Big),$$

каде $\chi^2_{n,\alpha/2}$ и $\chi^2_{n,1-\alpha/2}$ се такви да $P\{\chi^2_n > \chi^2_{n,\alpha/2}\} = \alpha/2$ и $P\{\chi^2_n > \chi^2_{n,1-\alpha/2}\} = 1-\alpha/2$ соодветно и χ^2_n е случајна променлива со хи-квадрат распределба со n степени на слобода. За дадените податоци имаме дека $n = 30$, а за веројатноста на доверба $1-\alpha = 0,9$, од каде $\alpha = 0,1$, па $\chi^2_{n,\alpha/2} = \chi^2_{30;0,05}$ и $\chi^2_{n,1-\alpha/2} = \chi^2_{30;0,95}$ се такви да $P\{\chi^2_{30} > \chi^2_{30;0,05}\} = 0,05$ и $P\{\chi^2_{30} > \chi^2_{30;0,95}\} = 0,95$ соодветно, од каде $\chi^2_{30;0,05} = 43,773$ и $\chi^2_{30;0,95} = 18,493$. За дадените податоци имаме $\sum_{i=1}^{30} X_i = 478$ и $\sum_{i=1}^{30} X_i^2 = 7970$, па за сумата во интервалот на доверба имаме

$$\sum_{i=1}^{30}(X_i - m)^2 = \sum_{i=1}^{30} X_i^2 - 2m\sum_{i=1}^{30} X_i + 30 \cdot m^2 = 7970 - 2 \cdot 16 \cdot 478 + 30 \cdot 16^2 = 354,$$

од каде бараниот 90% интервал на доверба за σ^2 е $I_{\sigma^2} = (8,08718; 19,1424)$.

Ирена Стојковска

Задача 4.46. Нека обележјето X на една популација има нормална $\mathcal{N}(m,\sigma^2)$ распределба. За податоците $x_1, x_2, ..., x_{15}$ кои одговараат на тоа обележје важи $\sum_{i=1}^{15} x_i = 15$ и $\sum_{i=1}^{15} x_i^2 = 27,3$. Врз основа на овие податоци најди 90% интервал на доверба за σ^2.

Решение. Обележјето X има $\mathcal{N}(m,\sigma^2)$ распределба, каде m и σ^2 се непознати параметри. Нека $(X_1, ..., X_n)$ е примерок кој одговара на обележјето X. Тогаш, интервал на доверба за дисперзијата σ^2 кога m не е познато, со веројатност на доверба $1-\alpha$ е

$$I_{\sigma^2} = \Big(\frac{n\overline{S}_n^2}{\chi^2_{n-1,\alpha/2}}, \frac{n\overline{S}_n^2}{\chi^2_{n-1,1-\alpha/2}}\Big),$$

каде $\overline{S}_n^2$ е дисперзијата на примерокот, а $\chi^2_{n-1,\alpha/2}$ и $\chi^2_{n-1,1-\alpha/2}$ се такви да $P\{\chi^2_{n-1} > \chi^2_{n-1,\alpha/2}\} = \alpha/2$ и $P\{\chi^2_{n-1} > \chi^2_{n-1,1-\alpha/2}\} = 1-\alpha/2$ соодветно и χ^2_{n-1} е случајна променлива со хи-квадрат распределба со $(n-1)$ степени на слобода. За дадените податоци $n = 15$, а за веројатноста на дверба $1-\alpha = 0,9$, од каде $\alpha = 0,1$ па $\chi^2_{n-1,\alpha/2} = \chi^2_{14;0,05}$ и $\chi^2_{n-1,1-\alpha/2} = \chi^2_{14;0,95}$ се такви да $P\{\chi^2_{14} > \chi^2_{14;0,05}\} = 0,05$ и $P\{\chi^2_{14} > \chi^2_{14;0,95}\} = 0,95$ соодветно, од каде $\chi^2_{14;0,05} = 23,685$ и $\chi^2_{14;0,95} = 6,571$. За броителот во интервалот на доверба имаме

$$\begin{aligned} n\overline{S}_n^2 &= 15\cdot\Big(\frac{1}{15}\sum_{i=1}^{15} X_i^2 - \Big(\frac{1}{15}\sum_{i=1}^{15} X_i\Big)^2\Big) = \sum_{i=1}^{15} X_i^2 - \frac{1}{15}\Big(\sum_{i=1}^{15} X_i\Big)^2 = \\ &= 27,3 - \frac{1}{15}\cdot 15^2 = 27,3 - 15 = 12,3, \end{aligned}$$

од каде бараниот 90% интервал на доверба за σ^2 е $I_{\sigma^2} = (0,519316; 1,87186)$.

Задача 4.47. При 7 расипувања на една машина измерени се следните броеви на непрекинато работење на машината во часови: 53, 48, 50, 54, 51, 50 и 51. Под претпоставка дека X - број на часови на непрекината работа на машината има експоненцијална распределба, најди 95% интервал на доверба за очекуваниот број на часови на непрекината работа на машината.

Решение. Обележјето X - број на часови на непрекината работа на машината има експоненцијална $\mathcal{E}(\beta)$ распределба, каде $\beta > 0$ е непознат параметар. Нека $(X_1, ..., X_n)$ е примерок кој одговара на обележјето X, тогаш статистиката $T = \frac{2}{\beta}\sum_{i=1}^{n} X_i \sim \chi^2_{2n}$ (претходно покажано) може да ја земеме за централна статистика при наоѓање на интервал на доверба за очекуваниот број часови на непрекината работа на машината $E(X) = \beta$. За да го одредиме $(1-\alpha)100\%$ интервал на доверба за β, бараме $a, b \in \mathbb{R}$ такви што $P\{a < T < b\} = 1-\alpha$, односно $P\{T \leq a\} + P\{T \geq b\} = \alpha$. Бараме интервал со што е можно помала должина, па земаме $P\{T \leq a\} = \alpha/2$ и $P\{T \geq b\} = \alpha/2$, од каде бидејќи $T \sim \chi^2_{2n}$ имаме дека $a = \chi^2_{2n,1-\alpha/2}$ и $b = \chi^2_{2n,\alpha/2}$. Сега,

$$\begin{aligned} 1-\alpha &= P\{a < T < b\} = P\{a < \frac{2}{\beta}\sum_{i=1}^{n} X_i < b\} = P\{\frac{1}{b} < \frac{\beta}{2\sum_{i=1}^{n} X_i} < \frac{1}{a}\} = \\ &= P\{\frac{2\sum_{i=1}^{n} X_i}{b} < \beta < \frac{2\sum_{i=1}^{n} X_i}{a}\} = P\{\frac{2n\overline{X}_n}{b} < \beta < \frac{2n\overline{X}_n}{a}\}, \end{aligned}$$

што значи $(1-\alpha)100\%$ интервал на доверба за β е интервалот

$$I_\beta = \Big(\frac{2n\overline{X}_n}{b}, \frac{2n\overline{X}_n}{a}\Big),$$

каде $a = \chi^2_{2n,1-\alpha/2}$ и $b = \chi^2_{2n,\alpha/2}$. Од $1-\alpha = 0,95$, следи $\alpha = 0,05$, потоа за дадените податоци имаме дека $n = 7$, од каде $a = \chi^2_{2n,1-\alpha/2} = \chi^2_{14;0,975} = 7,3578$ и $b = \chi^2_{2n,\alpha/2} = \chi^2_{14;0,025} = 26,3417$, а за броеителот во интервалот на доверба имаме

$$2n\overline{X}_n = 2\sum_{i=1}^{n} X_i = 2\cdot(53+48+50+54+51+50+51) = 714.$$

Значи, бараниот 95% интербвал на доверба за очекуваниот број часови на непрекината работа на машината е $I_\beta = (27,1053; 97,0399)$.

Задача 4.48. Врз основа на независни примероци $(X_1,...,X_{n_1})$ и $(Y_1,...,Y_{n_2})$ кои одговараат на обележјата $X \sim \mathcal{N}(m_1,\sigma_1^2)$ и $Y \sim \mathcal{N}(m_2,\sigma_2^2)$ соодветно, најди $100(1-\alpha)\%$ интервал на доверба за разликата $m_1 - m_2$.

Решение. Нека $X \sim \mathcal{N}(m_1,\sigma_1^2)$ и $Y \sim \mathcal{N}(m_2,\sigma_2^2)$, а $(X_1,...,X_{n_1})$ и $(Y_1,...,Y_{n_2})$ се примероци кои одговараат на обележјата X и Y соодветно. Тогаш,

$$T = \frac{\overline{X}_{n_1} - \overline{Y}_{n_2} - (m_1 - m_2)}{\sqrt{n_1\overline{S}_X^2 + n_2\overline{S}_Y^2}}\sqrt{\frac{n_1 n_2}{n_1+n_2}(n_1+n_2-2)} \sim t_{n_1+n_2-2},$$

(претходно покажано), каде $\overline{X}_{n_1}$ и $\overline{S}_X^2$ се средината и дисперзијата на примерокот $(X_1,...,X_{n_1})$, а $\overline{Y}_{n_2}$ и $\overline{S}_Y^2$ се средината и дисперзијата на примерокот $(Y_1,...,Y_{n_2})$, па статистиката T може да ја земеме за централна статистика за разликата $m_1 - m_2$. За да го одредиме $(1-\alpha)100\%$ интервал на доверба за $m_1 - m_2$, бараме $a, b \in \mathbb{R}$ такви што $P\{a < T < b\} = 1-\alpha$. Статистиката T има симетрична студентова $t_{n_1+n_2-2}$ распределба, па интервал со минимална должина се добива за $a = -b$, од каде последното равенство преминува во

$P\{|T|<b\}=1-\alpha$, односно $P\{|T|\geq b\}=\alpha$, од каде $b=t_{n_1+n_2-2,\alpha}$. Сега,

$$\begin{aligned}
&1-\alpha = P\{a<T<b\} = P\{-b<T<b\} = \\
= \ & P\{-b<\frac{\overline{X}_{n_1}-\overline{Y}_{n_2}-(m_1-m_2)}{\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}<b\} = \\
= \ & P\{\frac{-b\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}{\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}}<\overline{X}_{n_1}-\overline{Y}_{n_2}-(m_1-m_2)<\frac{b\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}{\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}}\} \\
= \ & P\{\overline{X}_{n_1}-\overline{Y}_{n_2}-\frac{b\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}{\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}}<m_1-m_2<\overline{X}_{n_1}-\overline{Y}_{n_2}+\frac{b\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}{\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}}\},
\end{aligned}$$

односно $(1-\alpha)100\%$ интервал на доверба за m_1-m_2 е интервалот

$$I=\Big(\overline{X}_{n_1}-\overline{Y}_{n_2}-\frac{b\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}{\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}},\overline{X}_{n_1}-\overline{Y}_{n_2}+\frac{b\sqrt{n_1\overline{S}_X^2+n_2\overline{S}_Y^2}}{\sqrt{\frac{n_1n_2}{n_1+n_2}(n_1+n_2-2)}}\Big),$$

каде $b=t_{n_1+n_2-2,\alpha}$.

Задачи за самостојна работа

Задача 4.49. При мерење на прирастот на масата (во грамови) кај примерок од 12 пилиња во текот на 24 саата, при примена на храна богата со протеини, добиени се следните резултати: 134, 146, 104, 119, 124, 161, 107, 83, 113, 129, 97, 123. Под претпоставка дека X - прираст на масата во грамови на едно пиле има нормална $\mathcal{N}(a,15^2)$ распределба, најди 95% интервал на доверба за очекуваниот прираст на масата на едно пиле.

Задача 4.50. За податоците $x_1,x_2,...,x_{10}$ кои одговараат на обележјето X со нормална распределба важи $\sum_{i=1}^{10}x_i=34$ и $\sum_{i=1}^{10}x_i^2=138$. За истото обележје извлечени се нови податоци $y_1,y_2,...,y_{15}$ за кои $\sum_{i=1}^{15}y_i=51$ и $\sum_{i=1}^{15}y_i^2=207$. За колку се променила должината на 90% интервал на доверба за математичкото очекување на обележјето X?

Задача 4.51. Обележјето X има $\mathcal{U}(0,2\theta)$ распределба, каде $\theta>0$ е непознат параметар. Нека $(X_1,...,X_n)$ е примерок кој одговара на обележјето X. Со помош на статистиката $\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ и со примена на централната гранична теорема, најди 90% интервал на доверба за θ, врз основа на податоците:

x_k	0,2	0,3	0,6	0,8	1,1	1,2	1,4
f_k	6	7	8	7	8	6	8

Ирена Стојковска

Потоа, врз основа на истите податоци, одреди точен интервал на доверба за θ од обликот $\left(\frac{X_{(n)}}{2}, \frac{X_{(n)}}{A}\right)$, каде $X_{(n)} = \max\{X_1, ..., X_n\}$ и спореди ги должините на добиените интервали на доверба.

Задача 4.52. Врз основа на независни примероци $(X_1, ..., X_{n_1})$ и $(Y_1, ..., Y_{n_2})$ кои одговараат на обележјата $X \sim \mathcal{N}(m_1, \sigma_1^2)$ и $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ соодветно, најди $100(1-\alpha)\%$ интервал на доверба за количникот σ_1^2/σ_2^2.
Упатство: За централна статистика земи ја

$$T = \frac{n_1(n_2-1)\sigma_2^2\overline{S}_X^2}{n_2(n_1-1)\sigma_1^2\overline{S}_Y^2} \sim F_{n_1-1,n_2-1}.$$