$$D(X) = \sum_{i=1}^{2}(x_i - p)^2 \cdot p_i = (1-p)^2 \cdot p + p^2 \cdot (1-p) = p \cdot (1-p).$$

Poznámka. Dá sa dokázať, že disperzia náhodnej premennej X s binomickým rozdelením pravdepodobnosti s parametrami $n \in N$, $p \in (0, 1)$ je $D(X) = n \cdot p$. Disperzia náhodnej premennej X s Poissonovým rozdelením $Po(\lambda)$ je $D(X) = \lambda$. Disperzia náhodnej premennej X s hypergeometrickým rozdelením s parametrami N, M, n je

$$D(X) = \frac{M}{N} \cdot n \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1}.$$

Definícia 3.14 Nech X je spojitá náhodná premenná s funkciou hustoty *f*. Disperzia $D(X)$ náhodnej premennej X je definovaná vzťahom

$$D(X) = \int_{-\infty}^{\infty} (x - E(X))^2 \cdot f(x)dx,$$

ak uvedený integrál konverguje.

Poznámka. Pri výpočte disperzie môžeme využiť nasledujúci upravený vzťah:

$$D(X) = E(X^2) - E^2(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - (\int_{-\infty}^{\infty} x \cdot f(x)dx)^2.$$

Príklad 3.19 Uvažujme náhodnú premennú X s rovnomerným rozdelením pravdepodobnosti na intervale (a, b). V príklade 3.17 sme vypočítali jej strednú hodnotu $E(X) = \dfrac{a+b}{2}$.

$$D(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - E^2(X) = \int_{a}^{b} x^2 \cdot \frac{1}{b-a} dx - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Poznámka. Dá sa dokázať, že disperzia náhodnej premennej X s normálnym rozdelením $N(\mu, \sigma^2)$ sa rovná parametru σ^2.

4 ZÁKLADNÝ A VÝBEROVÝ SÚBOR, NÁHODNÝ VÝBER

V popisnej štatistike bol základným pojmom pojem štatistický súbor. Bol to súbor štatistických jednotiek, na ktorých bol pozorovaný určitý štatistický znak. Súbor všetkých štatistických jednotiek, ktoré sú predmetom štatistického skúmania, budeme nazývať *základný súbor*. Málokedy však môžeme preskúmať celý základný súbor. Preto vlastnosti štatistických znakov na danom základnom súbore (najmä vtedy, ak je rozsiahly) skúmame tak, že z neho vyberieme určitý počet jednotiek, ktoré daný súbor reprezentujú. Túto množinu vybratých jednotiek voláme *výberový súbor*. Výberový súbor je akýmsi zmenšeným obrazom základného súboru. Jeho prvky musia byť zo základného súboru vybrané tak, aby sme dostali neskreslený obraz základného súboru.

Najpoužívanejším druhom výberového súboru je *náhodný výberový súbor*. Výberový súbor sa nazýva náhodný, ak o tom, či sa jednotka základného súboru vybrala alebo nie, rozhodla iba náhoda. Každý prvok základného súboru musí mať rovnakú pravdepodobnosť, že sa stane prvkom výberového súboru. Náhodnosť vybratia zabezpečuje, že výberový súbor je „verným“ obrazom základného súboru.

V praxi pri vytváraní výberového súboru musíme zabezpečiť realizáciu princípu náhodnosti. Najjednoduchšou technikou náhodného výberu je losovanie, často sa náhodný výber realizuje aj pomocou tabuliek náhodných čísel alebo pomocou náhodných čísel generovaných počítačom. Uvedieme príklad.

Príklad 4.1 Nech základným súborom je množina všetkých obyvateľov Slovenska starších ako 18 rokov. Určitá agentúra má formou ankety zistiť preferencie politických strán. Zrejme by bolo finančne a časovo veľmi náročné, keby v rámci prieskumu agentúra oslovila každého obyvateľa Slovenska staršieho ako 18 rokov. Z týchto dôvodov agentúra osloví iba určitý počet obyvateľov, ktorí budú tvoriť výberový súbor. Aby bol dodržaný princíp reprezentatívnosti, výber respondentov musí byť náhodný.

Cieľom štatistických analýz je dospieť na základe údajov o výberovom súbore ku všeobecným záverom o základnom súbore. Každý základný súbor je charakterizovaný určitými znakmi, ktoré môžeme opísať rozdelením pravdepodobnosti a číselnými charakteristikami, ktoré nazývame parametre znakov základného súboru. Parametrami znakov základného súboru sú napríklad stredná hodnota, rozptyl a korelačný koeficient. V reálnych situáciách nepoznáme rozdelenia pravdepodobnosti znakov na základnom súbore ani jeho parametre. Zo základného súboru urobíme náhodný výber, ktorý tvorí výberový štatistický súbor. Na základe hodnôt pozorovaného znaku na prvkoch výberového súboru robíme závery o rozdelení pravdepodobnosti znaku na základnom súbore resp. o jeho neznámych parametroch. Nedostávame presné výsledky, ale len odhady parametrov. Štatistická metóda, ktorá sa zaoberá úlohami získania čo najlepších odhadov neznámych parametrov, sa nazýva *teória odhadu*. Niekedy máme určité domnienky o rozdelení pravdepodobnosti znaku na základnom súbore alebo predpokladáme, že neznámy parameter sa rovná určitej hodnote. Overovanie správnosti týchto predpokladov nazývame *testovanie štatistických hypotéz*.

Skôr ako pristúpime k definovaniu ďalších pojmov, pre motiváciu uvedieme príklad.

Príklad 4.2 Predpokladajme, že treba vyhodnotiť pre potreby textilného priemyslu výšku desaťročných chlapcov Slovenskej republiky. Postupovalo by sa tak, že by sa náhodne vybralo povedzme 1000 desaťročných chlapcov. Množina týchto chlapcov (označme ju *M*) tvorí výberový súbor. Pozorovaným znakom *X* je výška chlapcov. Z matematického hľadiska je znak *X* reálnou funkciou, ktorá každému prvku $i \in M$ priradí číselnú hodnotu x_i. Teda X: $M \to R$, $X(i) = x_i$, $i = 1,2,...,1000$. Hodnoty x_i funkcie *X* sa menia pri každom prvku *i* dôsledkom množstva náhodných vplyvov. Pred meraním výšky chlapca nemôžeme presne predpovedať výsledok merania, určité hodnoty však môžeme vylúčiť (napríklad výšku pod 100 cm resp. nad 200 cm) a ďalšie hodnoty možno očakávať s väčšou alebo menšou pravdepodobnosťou. Hodnoty funkcie *X* závisia na náhode. Takéto funkcie sme v teórii pravdepodobnosti nazvali náhodné premenné. Každý kvantitatívny znak je z matematického hľadiska náhodná premenná.

Budeme predpokladať, že pozorovaný znak *X* má normálne rozdelenie $N(\mu, \sigma^2)$. Ak odmeriame výšku 1000 chlapcov z výberového súboru, dostaneme hodnoty $x_1, x_2, ..., x_{1000}$. Táto postupnosť je tzv. realizácia náhodného výberu $(X_1, X_2, ..., X_{1000})$, kde X_i je náhodná premenná označujúca výšku *i*- teho chlapca, $i = 1,2,...,1000$. Náhodné premenné X_i majú normálne rozdelenie, $X_i \sim N(\mu,\sigma^2)$, $i = 1,2,...,1000$. Keďže výška *i*- teho chlapca nezávisí od výšok ostatných chlapcov, náhodné premenné $X_1, X_2,..., X_{1000}$ sú navzájom nezávislé.

V nasledujúcej definícii zavedieme pojem *náhodný výber*, *realizácia náhodného výberu* a *výberová charakteristika*.

Definícia 4.1 Náhodným výberom s rozsahom n z rozdelenia s distribučnou funkciou F nazývame n-ticu navzájom nezávislých náhodných premenných $(X_1, X_2,..., X_n)$ s rovnakým pravdepodobnostným rozdelením s distribučnou funkciou F. n- ticu $(x_1, x_2,..., x_n)$ hodnôt náhodných premenných $X_1, X_2,..., X_n$ nazývame realizácia náhodného výberu. Funkcie náhodného výberu sa nazývajú výberové charakteristiky alebo štatistiky.

Poznámka. V predchádzajúcom príklade išlo o náhodný výber s rozsahom $n = 1000$ z normálneho rozdelenia $N(\mu,\sigma^2)$.

Príkladom výberovej charakteristiky je tzv. výberový aritmetický priemer $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

.

5 BODOVÝ A INTERVALOVÝ ODHAD

5.1 Bodový odhad

Predpokladajme, že pozorovaný znak X má pravdepodobnostné rozdelenie so strednou hodnotou μ a disperziou σ^2, pričom hodnoty týchto parametrov nepoznáme. V popisnej štatistike sme z nameraných údajov $x_1, x_2, ..., x_n$ vypočítali aritmetický priemer $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ a rozptyl $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$. Aritmetický priemer $\bar{x}$ je tzv. *bodovým odhadom* (presnejšie realizáciou bodového odhadu) strednej hodnoty μ. Bodovým odhadom disperzie σ^2 je výberový rozptyl $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$. Sú to takzvané bodové odhady preto, že sú vyjadrené jedným číslom. Okrem bodových odhadov sa používajú aj *intervalové odhady*, ktoré udávajú interval pokrývajúci odhadovaný parameter s určitou vopred predpísanou pravdepodobnosťou.

Stredná hodnota a disperzia základného súboru sú najdôležitejšie parametre základného súboru. Základný súbor však môže mať aj iné parametre. Neznámy parameter základného súboru budeme vo všeobecnosti označovať θ. Množinu všetkých hodnôt, ktoré môže parameter θ nadobúdať, nazývame *parametrický priestor*. Ak náhodná premenná X má diskrétne rozdelenie s neznámym parametrom θ, tak jej pravdepodobnostnú funkciu budeme označovať $p(x,\theta)$. Ak náhodná premenná X má spojité rozdelenie s neznámym parametrom θ, tak jej funkciu hustoty budeme označovať $f(x,\theta)$.

Definícia 5.1 Nech $(X_1, X_2, ..., X_n)$ je náhodný výber z rozdelenia s neznámym parametrom θ a nech $(x_1, x_2, ..., x_n)$ je jeho realizácia. Výberovým odhadom parametra θ nazývame štatistiku (náhodnú premennú) $T_n = T(X_1, X_2, ..., X_n)$. Bodový odhad parametra θ

dostaneme, ak z realizácie $(x_1, x_2, ..., x_n)$ vypočítame hodnotu realizácie výberového odhadu, t.j. hodnotu $t_n = T(x_1, x_2, ..., x_n)$.

Budeme definovať bodové odhady, ktoré majú „dobré vlastnosti". Najskôr budeme definovať tieto „dobré vlastnosti" odhadov.

Definícia 5.2 Nech $(X_1, X_2, ..., X_n)$ je náhodný výber zo základného súboru, ktorý má rozdelenie s neznámym parametrom θ. Výberový odhad $T_n = T(X_1, X_2, ..., X_n)$ nazývame nevychýleným odhadom parametra θ, ak $E(T_n) = \theta$. Ak $E(T_n) \neq \theta$, tak rozdiel $E(T_n) - \theta$ nazývame vychýlením odhadu T_n. Odhad T_n, pre ktorý platí $\lim_{n\to\infty} E(T_n) = \theta$, nazývame asymptoticky nevychýlený odhad parametra θ.

Príklad 5.1 Nech $(X_1, X_2, ..., X_n)$ je náhodný výber z rozdelenia s neznámou strednou hodnotou μ a neznámou disperziou σ^2. Potom sa dá dokázať, že výberový aritmetický priemer $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ je nevychýleným odhadom strednej hodnoty μ, rozptyl $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ je nevychýleným odhadom disperzie σ^2 a rozptyl $S_*^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ je asymptoticky nevychýleným odhadom disperzie σ^2. Preto pri odhade disperzie σ^2 budeme používať rozptyl S^2. Pri veľkom rozsahu výberového súboru možno používať aj rozptyl S_*^2. Pri praktickom výpočte z nameraných hodnôt $x_1, x_2, ..., x_n$ vypočítame ich aritmetický priemer $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ a hodnotu $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$. Pri veľkom rozsahu súboru môžeme počítať aj hodnotu $s_*^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$.

Poznámka. Je zrejmé, že možno definovať viacero nevychýlených odhadov parametra θ. Napríklad štatistika $T_n = \overline{X}$ aj štatistika $T_n = X_n$ sú nevychýlené odhady strednej hodnoty, pričom $D(\overline{X}) = \frac{\sigma^2}{n}$, $D(X_n) = \sigma^2$. Lepším odhadom je štatistika $\overline{X}$, lebo má menšiu disperziu.

Štatistika T_n je tzv. *efektívnym odhadom* parametra θ, ak T_n je nevychýleným odhadom parametra θ a má najmenšiu disperziu v triede všetkých nevychýlených odhadov parametra θ.

5.2 Intervalový odhad

Cieľom intervalového odhadu je nájsť taký (náhodný) interval, ktorý by hľadaný parameter pokrýval s dostatočne veľkou (vopred predpísanou) pravdepodobnosťou. V praxi je interval určený dvoma číslami, v pravdepodobnostnom modeli bude intervalový odhad určený dvoma štatistikami. V ďalšom budeme predpokladať, že základný súbor má normálne rozdelenie $N(\mu, \sigma^2)$ a budeme sa zaoberať intervalovými odhadmi jeho parametrov μ a σ^2.

5.2.1 Interval spoľahlivosti strednej hodnoty normálneho rozdelenia pri známej disperzii

Nech $(X_1, X_2, ..., X_n)$ je náhodný výber z rozdelenia $N(\mu, \sigma^2)$ s neznámou strednou hodnotou μ a známou disperziou σ^2. Výberovým odhadom strednej hodnoty μ je výberový aritmetický priemer $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, ktorý má rozdelenie $N(\mu, \frac{\sigma^2}{n})$. Dá sa dokázať, že štatistika

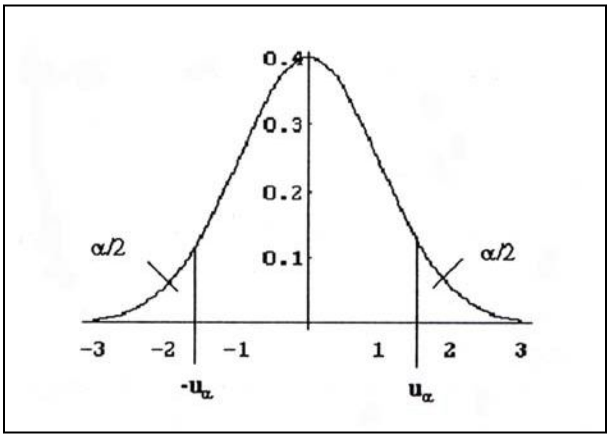
$$U = \frac{\overline{X} - \mu}{\sigma}\sqrt{n}$$

má rozdelenie $N(0, 1)$.

Pri intervalovom odhade a aj pri testovaní štatistických hypotéz používame pojem *kritická hodnota*. Nasledujúca definícia je definíciou kritickej hodnoty $N(0, 1)$ rozdelenia.

Definícia 5.2 Hodnotu u_α náhodnej premennej U s rozdelením $N(0, 1)$ nazývame kritickou hodnotou rozdelenia $N(0, 1)$, ak platí

$$P(|U| > u_\alpha) = \alpha .$$



Obr. 5.1.

Kritické hodnoty u_α sú pre dané hodnoty α uvedené v tabuľke č.12.2.
Je zrejmé, že platí rovnosť

$$P(|U| \le u_\alpha) = 1 - \alpha .$$

Pomocou tejto rovnosti odvodíme intervalový odhad parametra μ:

$$1 - \alpha = P(-u_\alpha \le U \le u_\alpha) = P(-u_\alpha \le \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \le u_\alpha) =$$

$$= P(\overline{X} - u_\alpha \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + u_\alpha \cdot \frac{\sigma}{\sqrt{n}}).$$

Dostali sme interval $\left\langle \overline{X} - u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}; \overline{X} + u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right\rangle$. Tomuto intervalu hovoríme $100 \cdot (1-\alpha)$ percentný *obojstranný interval spoľahlivosti*. Analogicky sa odvodia aj tzv. jednostranné intervaly spoľahlivosti.

Interval $\left\langle \overline{X} - u_{2\alpha} \cdot \frac{\sigma}{\sqrt{n}}; \infty \right)$ je $100 \cdot (1-\alpha)$ percentný ľavostranný interval spoľahlivosti pre strednú hodnotu μ.

Interval $\left(-\infty ; \overline{X} + u_{2\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right\rangle$ je $100 \cdot (1-\alpha)$ percentný pravostranný interval spoľahlivosti pre strednú hodnotu μ.

Získané výsledky interpretujeme nasledujúcim spôsobom: Ak v praktických situáciách vypočítame pomocou predchádzajúceho vzťahu pre zvolené α $100 \cdot (1-\alpha)$ percentné intervaly spoľahlivosti strednej hodnoty μ, tak vo veľkom počte prípadov v $1-\alpha$ percentách z nich bude vypočítaný interval strednú hodnotu μ naozaj pokrývať.

Príklad 5.2 95-percentným intervalom spoľahlivosti treba odhadnúť priemernú váhu vyrábaného výrobku. Preto sme náhodným výberom vybrali 100 výrobkov. Ich priemerná váha je 2,2 kg. Na základe dlhodobých pozorovaní vieme, že smerodajná odchýlka váhy výrobkov σ = 0,6 kg.

Riešenie. Základný súbor tvoria vyrábané výrobky, výberový súbor tvorí 100 náhodne vybraných výrobkov. Pozorovaným znakom X je váha výrobkov. Budeme predpokladať, že znak X má normálne rozdelenie $N(\mu, \sigma^2)$.

Keďže 1 - α = 0,95, α = 0,05. V tabuľke č. 12.2 vyhľadáme kritickú hodnotu $u_{0,05}$ = 1,96 a určíme požadovaný interval spoľahlivosti:

$$\left\langle \overline{X} - u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}; \overline{X} + u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right\rangle = \left\langle 2{,}2 - 1{,}96 \cdot \frac{0{,}6}{10}; 2{,}2 + 1{,}96 \cdot \frac{0{,}6}{10} \right\rangle = \langle 2{,}08; 2{,}32 \rangle.$$

95-percentným obojstranným intervalom spoľahlivosti pre priemernú váhu vyrábaného výrobku je interval $\langle 2{,}08;\ 2{,}32 \rangle$.

5.2.2 Interval spoľahlivosti strednej hodnoty normálneho rozdelenia pri neznámej disperzii

Situácia, keď poznáme disperziu σ^2, je zriedkavá. Častejšie musíme pracovať s odhadom neznámej disperzie σ^2, ktorým je štatistika $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Pre intervalový odhad strednej hodnoty μ normálneho rozdelenia v tomto prípade použijeme štatistiku

$$t = \frac{\overline{X} - \mu}{S}\sqrt{n},$$

ktorá má tzv. *Studentovo t- rozdelenie* o $n-1$ stupňoch voľnosti. Namiesto kritických hodnôt u_α normálneho rozdelenia použijeme kritické hodnoty Studentovho *t*- rozdelenia, ktoré sú uvedené v tabuľke č. 12.3 a sú definované nasledovne.

Definícia 5.3 Nech náhodná premenná X má Studentovo *t*- rozdelenie o n stupňoch voľnosti. Kritickou hodnotou Studentovho *t*- rozdelenia nazývame hodnotu $t_\alpha(n)$, pre ktorú platí

$$P(|X| > t_\alpha(n)) = \alpha .$$

Intervaly spoľahlivosti pre strednú hodnotu normálneho rozdelenia pri neznámom rozptyle sa odvodia rovnako ako pri známom rozptyle.

Interval $\left\langle \overline{X} - t_\alpha(n-1)\cdot\frac{S}{\sqrt{n}};\ \overline{X} + t_\alpha(n-1)\cdot\frac{S}{\sqrt{n}} \right\rangle$ je $100\cdot(1-\alpha)$ percentný obojstranný interval spoľahlivosti pre strednú hodnotu μ, ak σ^2 nepoznáme.

Interval $\left\langle \overline{X} - t_{2\alpha}(n-1)\cdot\frac{S}{\sqrt{n}};\ \infty \right)$ je $100\cdot(1-\alpha)$ percentný ľavostranný interval spoľahlivosti pre strednú hodnotu μ.

Interval $\left(-\infty;\ \overline{X} + t_{2\alpha}(n-1)\cdot\frac{S}{\sqrt{n}} \right\rangle$ je $100\cdot(1-\alpha)$ percentný pravostranný interval spoľahlivosti pre strednú hodnotu μ.

Príklad 5.3 Treba odhadnúť 95-percentným intervalovým odhadom priemernú spotrebu oleja pre náterové hmoty. Spotrebu oleja sme namerali na n = 12 vzorkách, pričom sme vypočítali aritmetický priemer spotreby $\bar{x}$ = 14,306 a rozptyl spotreby s^2 = 0,327.

Riešenie. Pozorovaným znakom X je spotreba oleja. Budeme predpokladať, že znak X má normálne rozdelenie $N(\mu,\sigma^2)$, μ a σ^2 sú neznáme parametre. Pre α = 0,05 nájdeme v tabuľke č.12.3 kritickú hodnotu $t_{0,05}(11)$ = 2,2. Určíme požadovaný interval spoľahlivosti:

$$\left\langle \overline{X} - t_{\alpha}(n-1)\cdot\frac{S}{\sqrt{n}};\ \overline{X} + t_{\alpha}(n-1)\cdot\frac{S}{\sqrt{n}} \right\rangle = \left\langle 14{,}306 - 2{,}2\cdot\frac{0{,}57}{\sqrt{12}}; 14{,}306 + 2{,}2\cdot\frac{0{,}57}{\sqrt{12}} \right\rangle =$$
$$= \langle 13{,}938;\ 14{,}662 \rangle.$$

95-percentným intervalovým odhadom priemernej spotreby oleja pre náterové hmoty je interval $\langle 13{,}938;\ 14{,}662 \rangle$.

5.2.3 Interval spoľahlivosti pre disperziu normálneho rozdelenia

Nech $(X_1, X_2, ..., X_n)$ je náhodný výber z rozdelenia $N(\mu,\sigma^2)$. Pre odvodenie intervalu spoľahlivosti pre neznámy parameter σ^2 použijeme štatistiku

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2},$$

ktorá má χ^2 - rozdelenie s n -1 stupňami voľnosti. Kritické hodnoty tohto rozdelenia sú uvedené v tabuľke č. 12.4 a sú definované nasledovne.

Definícia 5.4 Nech náhodná premenná X má χ^2 - rozdelenie s n stupňami voľnosti. Kritickou hodnotou χ^2 - rozdelenia sa nazýva hodnota $\chi^2_{\alpha}(n)$ náhodnej premennej X, pre ktorú platí

$$P(X > \chi^2_{\alpha}(n)) = \alpha.$$

Interval $\left\langle \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}; \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \right\rangle$ je $100\cdot(1-\alpha)$ percentný obojstranný interval spoľahlivosti pre disperziu σ^2.

Interval $\left\langle \frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}; \infty \right)$ je $100\cdot(1-\alpha)$ percentný ľavostranný interval spoľahlivosti pre disperziu σ^2.

Interval $\left(0\,; \frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)} \right\rangle$ je $100\cdot(1-\alpha)$ percentný pravostranný interval spoľahlivosti pre disperziu σ^2.

Príklad 5.4 Pre určenie poistnej zásoby určitého druhu tovaru je potrebné odhadnúť 95-percentným intervalovým odhadom rozptyl času, za aký dodávateľ vybaví objednávku. Náhodne bolo vybratých 20 objednávok, u ktorých sa zaznamenala dĺžka ich vybavenia v dňoch. Boli získané tieto údaje: 18, 12, 18, 13, 13, 16, 11, 17, 20, 14, 15, 14, 16, 15, 15, 17, 13, 14, 16, 13.

Riešenie. Budeme predpokladať, že čas vybavenia objednávky je náhodná premenná s normálnym rozdelením $N(\mu,\sigma^2)$. Intervalovým odhadom budeme odhadovať parameter σ^2 tohto rozdelenia. Pre určenie intervalu spoľahlivosti vypočítame

$$(n-1)\cdot s^2 = \sum_{i=1}^{20}(x_i-\overline{x})^2 = 98.$$

Keďže 1 - α = 0,95, α = 0,05. V tabuľke č. 12.4 nájdeme príslušné kritické hodnoty: $\chi^2_{1-\frac{\alpha}{2}}(19) = \chi^2_{0,975}(19) = 8,91$; $\chi^2_{\frac{\alpha}{2}}(19) = \chi^2_{0,025}(19) = 32,9$.

Dostávame interval: $\left\langle \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}; \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \right\rangle = \left\langle \frac{98}{32,9}; \frac{98}{8,91} \right\rangle = \langle 3; 11 \rangle$.

Rozptyl času vybavenia objednávky dodávateľom pokrýva s 95-percentnou spoľahlivosťou interval $\langle 3; 11 \rangle$.

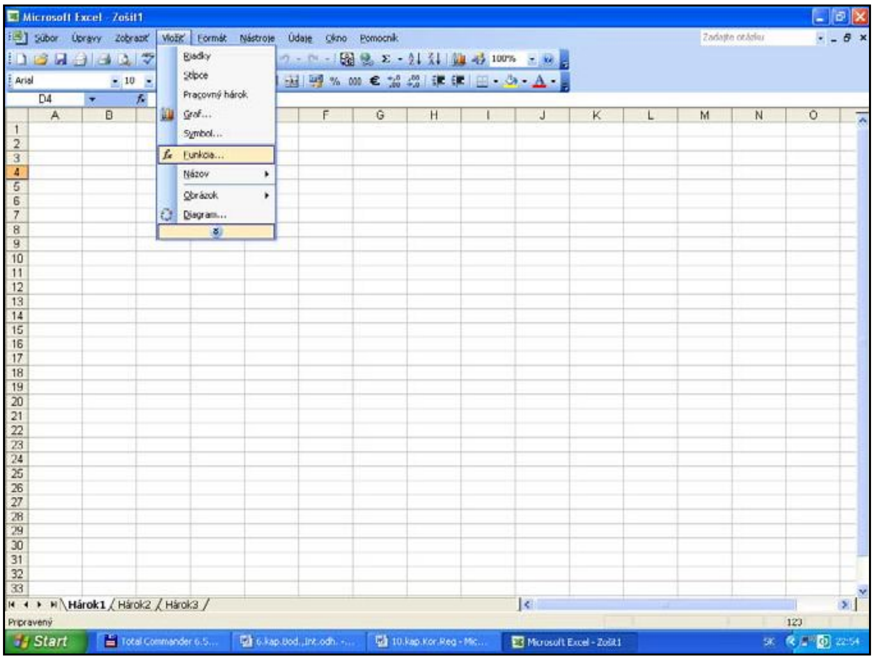
5.3 Intervaly spoľahlivosti v programe Excel

Pomocou programu Excel sa dá určiť interval spoľahlivosti iba pre strednú hodnotu pri známom rozptyle.

Postup práce pri určovaní intervalu spoľahlivosti pomocou programu Excel vysvetlíme na príklade 5.2, ktorý sme riešili v časti 5.2.1.

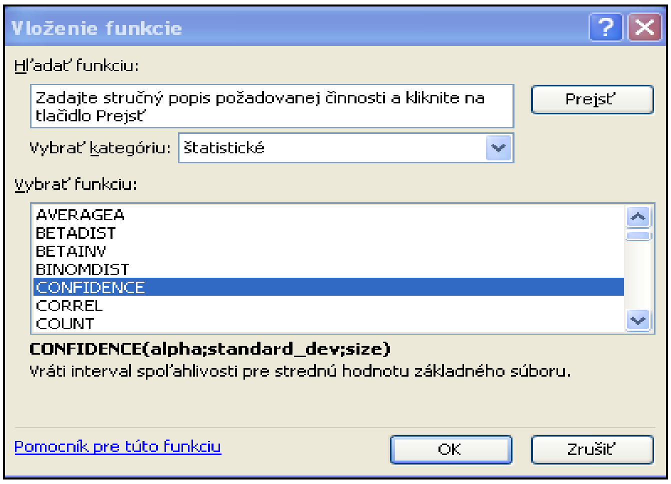
Zopakujeme jeho zadanie**:** 95-percentným intervalom spoľahlivosti treba odhadnúť priemernú váhu vyrábaného výrobku. Preto sme náhodným výberom vybrali 100 výrobkov. Ich priemerná váha je 2,2 kg. Na základe dlhodobých pozorovaní vieme, že smerodajná odchýlka váhy výrobkov σ = 0,6 kg.

Budeme postupovať nasledovne. V ponuke programu Excel postupne potvrdíme príkazy ***Vložiť – funkcia*** (obr. 5.2).



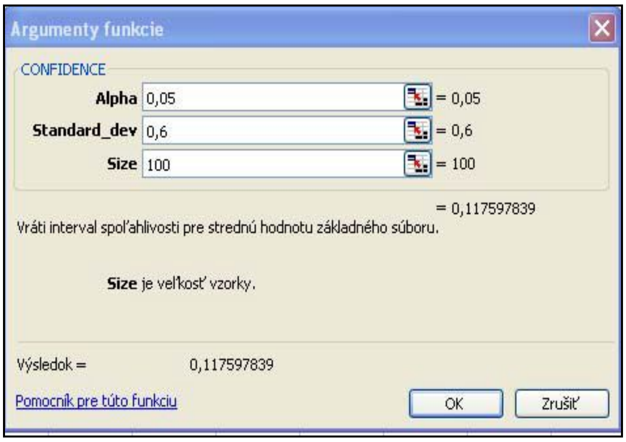
Obr. 5.2.

Po potvrdení ponuky sa zobrazí tabuľka s názvom ***Vloženie funkcie*** (obr. 5.3), kde v okienku ***Vybrať kategóriu*** zvolíme „štatistické“ a v kolónke ***Vybrať funkciu*** označíme CONFIDENCE .



Obr. 5.3.

Po potvrdení tlačidlom ***OK*** dostaneme tabuľku na obrázku 5.4.



Obr. 5.4.

V tabuľke s názvom ***Argumenty funkcie*** v okienku ***Alpha*** zvolíme hladinu významnosti α. Zvolili sme α = 0,05. Do okienka

Standard dev zapíšeme hodnotu smerodajnej odchýlky σ = 0,6 a do kolónky ***Size*** rozsah súboru. V našom prípade je rozsah súboru rovný 100.

Po potvrdení tlačidlom ***OK*** dostaneme hodnotu Δ tzv. prípustnej chyby odhadu, ktorá je definovaná vzťahom $\Delta = u_\alpha \cdot \dfrac{\sigma}{\sqrt{n}}$.

Do intervalu

$$\left\langle \overline{X} - u_\alpha \cdot \frac{\sigma}{\sqrt{n}};\ \overline{X} + u_\alpha \cdot \frac{\sigma}{\sqrt{n}} \right\rangle = \left\langle \overline{X} - \Delta;\ \overline{X} + \Delta \right\rangle$$

dosadíme vypočítanú hodnotu prípustnej chyby odhadu Δ = 0,11759 a hodnotu aritmetického priemeru $\bar{x}$ = 2,2. Pre zvolenú hladinu významnosti α = 0,05 dostaneme nasledujúci 95-percentný obojstranný interval spoľahlivosti:

$$\langle 2{,}2 - 0{,}11759\,;\ 2{,}2 + 0{,}11759 \rangle = \langle 2{,}08\,;\ 2{,}317 \rangle\,.$$