

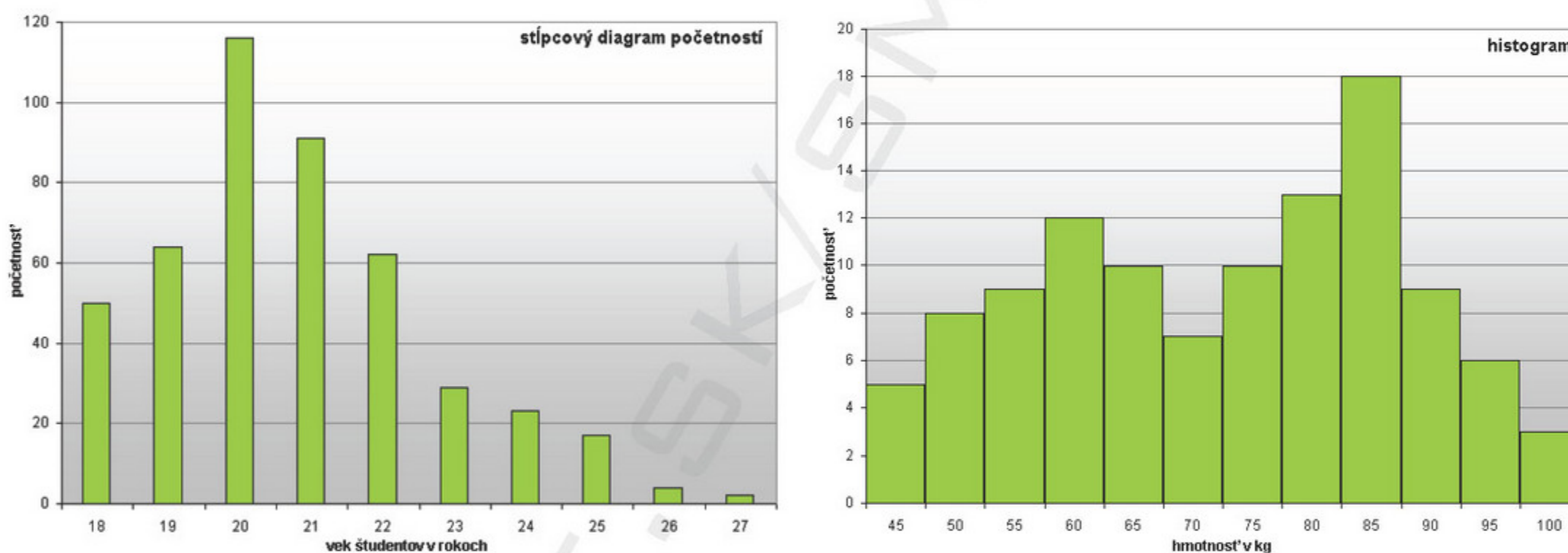
Základné štatistické pojmy sú obsiahnuté v norme v STN ISO 3534-1 (01 0216) [42].

Všetky skúmané štatistické jednotky tvoria tzv. základný súbor. Tie štatistické jednotky, ktoré sa zo skúmaného základného súboru vybrali a na základe ktorých sa robí o ňom úsudok, tvoria tzv. výberový súbor. Úsudky, ktoré sa urobia o príslušnom súbore na základe vybraných vzoriek, sú založené na skúsenosti, že vybrané vzorky viac alebo menej presne charakterizujú súbor, z ktorého sú vybrané.

V tejto kapitole nie sú definície vzťahované na základný súbor meraní ale len na množinu pozorovaní alebo výber zo základného súboru.

Stĺpcový diagram početností je diagram, kde pre jednotlivú hodnotu x_i (napríklad vek študentov v rokoch) skupiny meraných objektov (pre n študentov) je pridelená početnosť výskytu tejto hodnoty. Pre namerané hodnoty je možné vytvoriť tabuľku početností a stĺpcový diagram početností (Obr.3.1 vľavo).

Histogram: V prípade, že sa rovnaké merané hodnoty nachádzajú v súbore väčšinou v malom počte (napr. pri zisťovaní váhy študentov na stotiny kilogramu), stĺpcový diagram nemá zmysel. Tabuľka početností by bola neúmerne rozsiahla a početnosti by boli takmer všetky rovné 1. V takom prípade sa hodnoty zatriedia do tried (intervalov) rovnakej šírky tak, aby pokrývali všetky namerané hodnoty a vo väčšine tried bolo väčšie množstvo hodnôt. Pre každú triedu sa zvolí jedna jeho reprezentatívna hodnota (najčastejšie stred intervalu) ako tzv. triedna hodnota. Histogram je teda diagram, v ktorom každej triede, čiže intervalu, prislúcha stĺpec, ktorého šírka je rovná šírke triedy a jeho výška je rovná početnosti hodnôt v triede (Obr.3.1 vpravo).



Obr.3.1 Stĺpcový diagram a histogram

Charakteristiky polohy definujú priemernú alebo orientačnú hodnotu, okolo ktorej sa sústreďujú všetky merané hodnoty.

Priemery: Najbežnejším vyjadrením miery polohy skúmanej veličiny vo výberovom súbore je aritmetický (výberový) priemer označovaný \bar{x} . Je definovaný ako súčet hodnôt delený ich počtom

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.1)$$

Medzi iné odhady strednej hodnoty základného súboru patria napríklad aj geometrický alebo harmonický priemer, výberový medián a výberový modus.

Modus \hat{x} je najčastejšie sa vyskytujúca hodnota. Je to hodnota s najväčšou početnosťou.

Medián \tilde{x} konečnej množiny hodnôt je tá hodnota, ktorá rozdeľuje množinu na dve rovnaké časti, tak, že počet hodnôt rovných alebo väčších ako medián sa rovná počtu hodnôt menších alebo rovných ako medián.

Základným predpokladom pre určenie mediánu je *vzostupné (alebo zostupné) usporiadanie súboru hodnôt*. V prípade ak je počet hodnôt n číslo nepárne, je mediánom hodnota nachádzajúca sa v strede tohto usporiadaného súboru

$$\tilde{x} = x_{((n+1)/2)} \quad (3.2)$$

Ak je n číslo párne, je mediánom priemerná hodnota z dvoch prostredných hodnôt $x_{(n/2)}$, $x_{(n/2+1)}$ a vypočíta sa ako aritmetický priemer týchto dvoch hodnôt

$$\tilde{x} = \frac{x_{(n/2)} + x_{(n/2+1)}}{2} \quad (3.3)$$

Kvantily sú obecné také hodnoty, ktoré delia všetky hodnoty veličiny X usporiadané podľa veľkosti na niekoľko rovnako veľkých častí. Najbežnejšie používaný je 50%-ný kvantil, tzv. medián, ktorý sa okrem označenia \tilde{x} značí aj $x_{0.50}$. Často sa využívajú tiež kvartily (delia súbor na 4 rovnaké časti), decily (delia súbor na 10 rovnakých častí) a percentily (delia súbor na 100 rovnakých častí).

Kvartily rozdeľujú uvedený súbor na 4 približne rovnako početné časti. Sú tri: dolný kvantil $x_{0.25}$, medián $x_{0.50}$ a horný kvantil $x_{0.75}$.

Charakteristiky variability definujú meranú veličinu v celom súbore objektov jediným číslom, ktorým je možné usúdiť, ako veľmi sú tieto hodnoty v súbore rozptýlené. To znamená, ako sú vzdialené od priemernej hodnoty veličiny súboru.

Rozptyl je najpoužívanejšou mierou variability veličiny v danom súbore nameraných hodnôt. Určí sa zo vzťahu

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.4)$$

Smerodajná odchýlka je miera variability, ktorá vychádza z rozptylu. Je uvedená ako druhá odmocnina z hodnoty rozptylu a udáva sa v rovnakých jednotkách ako meraná veličina. Ak je meraná hodnota v mm, smerodajná odchýlka je tiež v mm a rozptyl je v mm^2 .

V laboratórnych podmienkach pracujeme s menším počtom hodnôt (výberovým súborom). Zo štatistického hľadiska ide potom o smerodajnú odchýlku z výberového súboru označovanú s . Pre výberové súbory s počtom hodnôt menším ako 30 sa smerodajná odchýlka vypočítava zo vzťahu

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.5)$$

Smerodajná odchýlka je zhruba definovaná ako priemerná odchýlka hodnôt od aritmetického priemeru.

Variačné rozpätie je rozdiel najväčšej (maximálnej) a najmenej (minimálnej) hodnoty v danom súbore meraných hodnôt

$$R = x_{\max} - x_{\min} \quad (3.6)$$

Kvôli vysokej citlivosti na extrémne odľahlé hodnoty (hrubé chyby), nie je variačné rozpätie často vhodným ukazovateľom variability.

Medzikvantilové rozpätie je vhodnejším ukazovateľom variability ako variačné rozpätie.

Jedným z medzikvantilových rozpätí je napríklad aj medzikvantilové rozpätie, ktoré je určené rozdielom horného (75%-ného) a dolného (25%-ného) kvartilu

$$R_q = x_{0.75} - x_{0.25} \quad (3.7)$$

Pre výpočet najväčšej (maximálnej) pravdepodobnej hodnoty platí

$$x_{P_{\max}} = \bar{x} + k \cdot s \quad (3.8)$$

Pre výpočet najmenej (minimálnej) pravdepodobnej hodnoty platí

$$x_{P_{\min}} = \bar{x} - k \cdot s \quad (3.9)$$

kde k je koeficient rozšírenia daný rozdelením pravdepodobnosti výskytu odchýlok.

Pri odhade výslednej hodnoty merania je potrebné zistiť, ako sú jednotlivé merané hodnoty rozptýlené okolo strednej hodnoty.

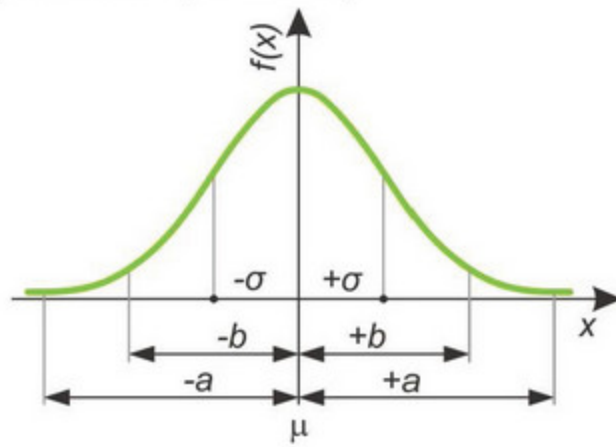
Ak nie je možné zodpovedne rozhodnúť o rozložení pravdepodobnosti výskytu odchýlok a je možné predpokladať, že hodnoty sa môžu vyskytovať v rozsahu nameraných hodnôt s rovnakou pravdepodobnosťou, potom volíme rovnorné rozdelenie. Tento prípad je najjednoduchší a používa sa najčastejšie.

Ak pravdepodobnosť výskytu odchýlok s ich rastúcou hodnotou klesá a najväčšiu pravdepodobnosť majú odchýlky malé, je vhodnou aproximáciou normálne (Gaussovo) alebo trojuholníkové (Simpsonovo) rozdelenie.

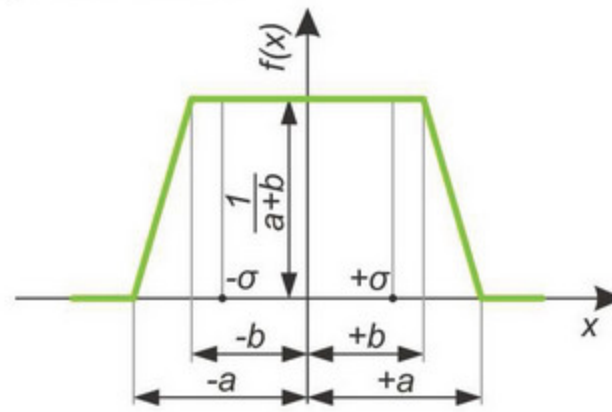
Normálne rozdelenie sa často nazýva aj zákon chýb, nakoľko je vhodnou aproximáciou popisujúcou správanie sa náhodných chýb pri meraní.

Najčastejšie typy rozdelenia pravdepodobnosti aj s príslušnými parametrami je možné vidieť na Obr.3.2.

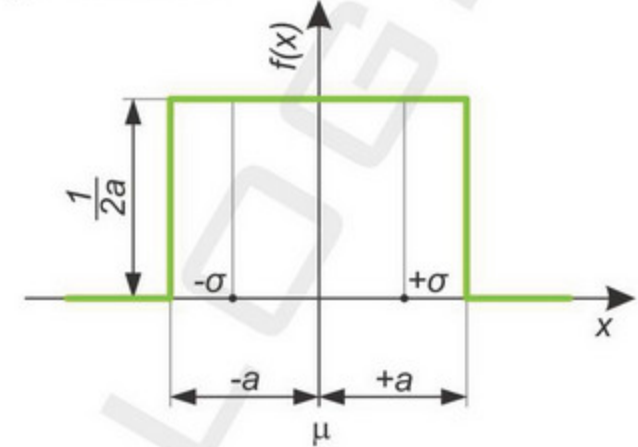
a) normálne (Gaussovo)



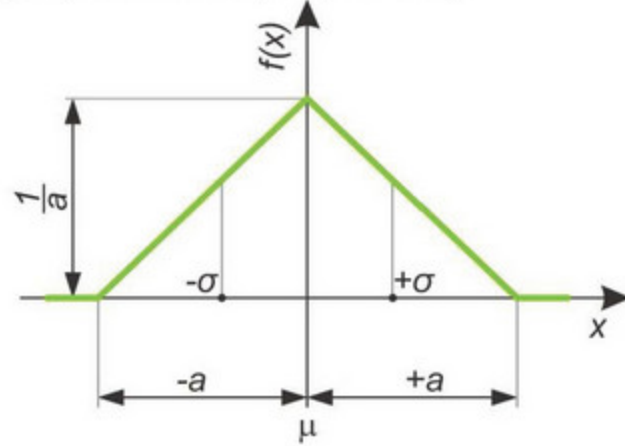
b) lichobežníkové



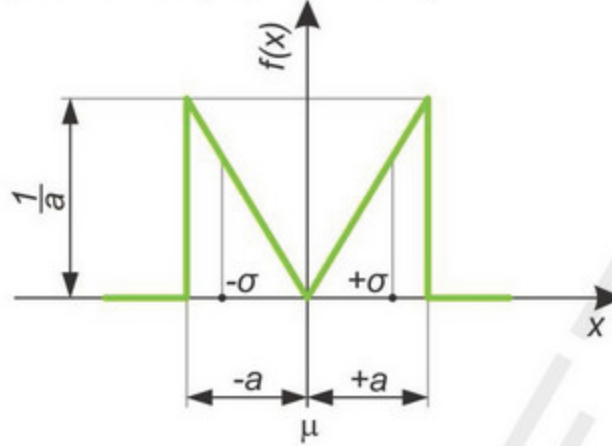
c) rovnomerné



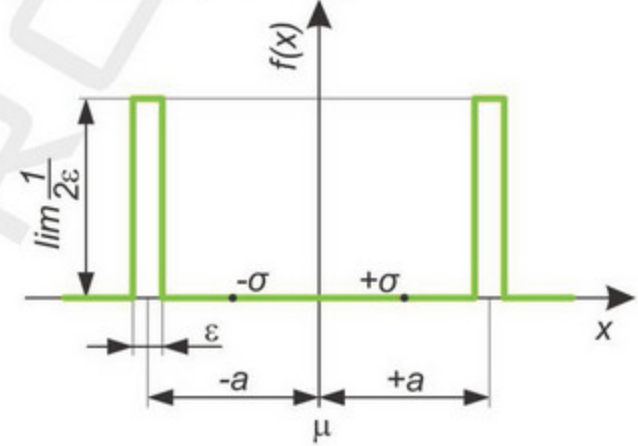
d) trojuholníkové (Simpsonovo)



e) bimodálne (trojuholníkové)



f) bimodálne (Diracovo)



Obr.3.2 Najčastejšie typy rozdelenia pravdepodobnosti.

Koeficienty rozšírenia: a) $a=3, b=2$; b) $a=2,32$ (pri $b=a/3$), $a=2,19$ (pri $b=a/2$), $a=2,04$ (pri $b=2a/3$)

c) $a=\sqrt{3}$ d) $a=\sqrt{6}$ e) $a=\sqrt{2}$ f) $a=1$

ÚLOHA Č.3

DIGITÁLNE PRACOVNÉ MERADLÁ - ZBER DÁT

Digitálne pracovné meradlá

Od osemdesiatych rokov sa začali používať digitálne posuvné meradlá (Obr.3.3) aj mikrometre (Obr.3.4), kde nónius a pomocnú stupnicu nahradili inkrementálny snímač a číslcový displej. Novou konštrukciou a princípom odmeriavania sa zvýšila presnosť týchto meradiel. Digitálne posuvné meradlá štandardne odmeriavajú s číslcovým krokom 0,01 mm a digitálne mikrometre s číslcovým krokom 0,001 mm.

Vďaka číslcovému zobrazovaniu je údaj jednoducho a rýchlo odčítateľný a eliminuje sa tým aj chyba odčítania hodnoty pri zlom uhle pohľadu (tzv. paralaxná chyba). Meradlá sú vďaka mikroprocesorovej technike vybavené dodatočnými funkciami ako nulovanie stupnice v ľubovoľnej polohe, prepínanie medzi zobrazením hodnoty v milimetroch alebo palcoch, zapamätanie si nameranej hodnoty aj pri ďalšom pohybe čelustí a možnosť komunikácie s nadradeným meracím systémom (osobný počítač, štatistické počítačové systémy, atď.), ktorý môže slúžiť aj ako vyhodnocovacia jednotka.



Obr.3.3 Popis základných častí digitálneho posuvného meradla



Obr.3.4 Popis základných častí digitálneho mikrometra

Zber dát pre štatistické vyhodnotenie merania

Digitálne pracovné meradlá sa pripájajú k PC stanici cez prevodník, ktorý umožňuje prenos dát z meradiel s rozhraním DIGIMATIC do počítača so sériovým rozhraním RS232 (Obr.3.5). Pre pripojenie jedného meradla sa využíva prevodník DMX-1, pre pripojenie až 3 meradiel je možné využiť trojkanálový prevodník DMX-3. Merané dáta sú spracovávané štatistickým softvérom od firmy Q-DAS v module Procella. Priamym prepojením a on-line zberom dát je možné rýchlo vyhodnotiť základné štatistické ukazovatele s možnou vizuálnou interpretáciou výsledkov vo forme grafov.



Obr.3.5 Zapojenie digitálnych pracovných meradiel

Meradlá je možné prepojiť aj so špeciálnymi štatistickými počítačmi (Obr.3.6), ktoré dokážu namerané hodnoty uchovávať v internej pamäti a neskôr z nich vypočítať základné parametre. V malom prístroji sa nachádza aj tlačiareň pre rýchlu tlač výsledkov na blok. Dáta je možné samozrejme aj ukladať pre neskoršie spracovanie na PC. Vďaka akumulátorom sú tieto prístroje prenosné.

Jednoduchšie prevedenia ponúkajú vyhodnotenie základných štatistických údajov ako napr. počet meraní, minimálna a maximálna hodnota, rozptyl hodnôt, priemerná hodnota, smerodajná odchýlka atď. Novšie prístroje dokonca vytlačia histogram alebo graf priebehu meraní. Vyššie verzie ponúkajú mnoho ďalších funkcií pre štatistickú reguláciu procesov (SPC).



Obr.3.6 Štatistické počítačadlo od firmy Mitutoyo