

ŠTATISTIKA V RADIČNEJ FYZIKE

OBSAH

1	Merania, neistoty a korelácie	3
1.1	Popis dát	3
1.1.1	Typy dát	4
1.1.2	Zobrazovanie dát	5
1.2	Priemery	6
1.2.1	Aritmetický priemer	6
1.2.2	Alternatívy aritmetickému priemeru	7
1.3	Meranie rozmazania	8
1.3.1	Rozptyl	8
1.3.2	Štandardná odchýlka	9
1.3.3	Alternatívne merania rozmazania	12
1.4	Vlastnosti kvantitatívnych dát	12
1.4.1	Miery polohy (central tendency - location)	12
1.4.2	Miery variability (variation - dispersion)	12
1.4.3	Miery tvaru (shape) (šikmosť, špicatosť);	13
1.5	Kovariancia a korelácia	14
1.5.1	Kovariancia	14
1.5.2	Korelácia	15
1.5.3	Viac než dve premenné	16
	Kontrolné otázky:	17
	Úlohy:	17
	Súhrn	18
2	Teoretické rozdelenia	19
2.1	Pravdepodobnosť	19
2.1.1	Zákon veľkých čísel	20
2.2	Očakávané hodnoty	21
2.3	Rozdelenie hustoty pravdepodobnosti	21
2.3.1	Binomické rozdelenie	22
2.3.2	Poissonove rozdelenie	24
2.3.3	Dve Poissonove rozdelenia	26
2.3.4	Gaussovské rozdelenie	26
2.3.5	Binormálne rozdelenie	30
2.4	Iné rozdelenia	31
2.4.1	Rovnomerné rozdelenie	31
2.4.2	Weibullove rozdelenie	32
2.4.3	Breit-Wignerove alebo Cauchyho rozdelenie	32
	Kontrolné otázky	33
	Úlohy	33
	Súhrn	33
3	Neistoty (Errors)	35
3.1	Úvod	35

3.2	Prečo sú neistoty gaussovské?.....	36
3.2.1	Centrálna limitná veta (central limit theorem - CLT)	36
3.3	Práca s neistotami.....	37
3.3.1	Opakovanie meraní	37
3.3.2	Priemerovanie vážených meraní	38
3.4	Kombinácie neistôt.....	38
3.4.1	Jedna premenná.....	38
3.4.2	Funkcia dvoch a viacerých premenných.....	39
3.4.3	Pravidlá kombinácie neistôt	40
3.4.4	Relatívne neistoty	40
3.4.5	Niekoľko funkcií niekoľkých premenných.....	41
3.4.6	Systematické neistoty	42
	Kontrolné otázky.....	43
	Úlohy.....	43
	Súhrn	44
4	Odhady (Estimation).....	45
4.1	Estimátory	45
4.1.1	Funkcia vierohodnosti.....	46
4.1.2	Niektoré základné estimátory	47
4.2	Maximálna vierohodnosť (likelihood).....	49
4.2.1	ML : Konzistentnosť, vychýlenosť a invariancia.....	51
4.2.2	Maximálna vierohodnosť pri veľkých N.....	52
4.2.3	Neistoty v ML estimátoroch	52
4.2.4	Viaceré premenné	53
4.2.5	Rozšírená maximálna vierohodnosť	54
4.3	Metóda momentov	55
4.4	Maximálna vierohodnosť a najmenšie štvorce	55
4.5	Stratifikované vzorkovanie	56
	kontrolné otázky	57
	Úlohy.....	57
	Súhrn	58
5	Metóda najmenších štvorcov	59
5.1	jednoduchá úmernosť	60
5.2	Lineárny fit	60
5.3	Váhovaný lineárny fit.....	61
5.4	Systematické chyby a lineárny fit	62
5.5	Fitovanie binovaných dát	62
5.6	χ^2 rozdelenie	63
5.7	Chyby u x aj y	65
5.8	Nelineárna metóda najmenších štvorcov	66
	Úlohy.....	66
	Kontrolné otázky.....	67
	Súhrn	67
6	Pravdepodobnosť a spoľahlivosť (Probability and Confidence).....	69
6.1	Pravdepodobnosť.....	70
6.1.1	Matematická pravdepodobnosť	70
6.1.2	Empirická pravdepodobnosť.....	70
6.1.3	3. Objektívna - sklon.....	70
6.1.4	Subjektívna pravdepodobnosť (Bayesova štatistika).....	71
6.2	Úroveň spoľahlivosti (Confidence level).....	72

6.2.1 Úroveň spoľahlivosti v popisnej štatistike.....	72
6.2.2 Intervaly spoľahlivosti v odhadoch.....	73
6.2.3 Úrovne spoľahlivosti z Gaussovho rozdelenia.....	74
6.2.4 Binomický interval spoľahlivosti.....	75
6.2.5 Poissonovské intervaly spoľahlivosti.....	76
6.2.6 Studentove t rozdelenie.....	76
Úlohy.....	79
Kontrolné otázky.....	79
Súhrn.....	80

Pri spracovaní a vyhodnocovaní väčšieho množstva údajov je nevyhnutné použiť metódy štatistiky, pričom pôvod a obsah dát môže byť veľmi rôznorodý. Štatistika je nástroj, ktorý pomáha plánovať a realizovať experimenty, analyzovať a interpretovať získané výsledky. Pre radiačnú fyziku sú typické hromadné údaje takže v tejto vednej oblasti nie je možné vyhnúť sa využitiu štatistických metód. Aby sme tento nástroj vedeli používať efektívne, je potrebné poznať ako pracuje, aké sú jeho možnosti a kedy ktoré jeho súčasti využívať. Štatistiku môžeme rozdeliť na popisnú, ktorá sa zaoberá získavaním, charakterom a prezentáciou dát a matematickú, ktorá sa zaoberá odhadmi a testovaním hypotéz. Cieľom popisnej štatistiky je popis dát, zatiaľ čo cieľom matematickej štatistiky je rozhodnúť o charakteristikách základného súboru dát.

1 MERANIA, NEISTOTY A KORELÁCIE

UČEBNÉ CIELE

Študent by sa mal naučiť popísať a charakterizovať súbor kvantitatívnych dát, optimálne ho zobraziť, mal by sa naučiť vypočítať priemer, štandardnú odchýlku, medián, modus, šikmosť, špicatosť, vedieť určiť či súbory dát sú vo vzájomnom vzťahu alebo sú vzájomne na sebe nezávislé.

KLÚČOVÉ SLOVÁ

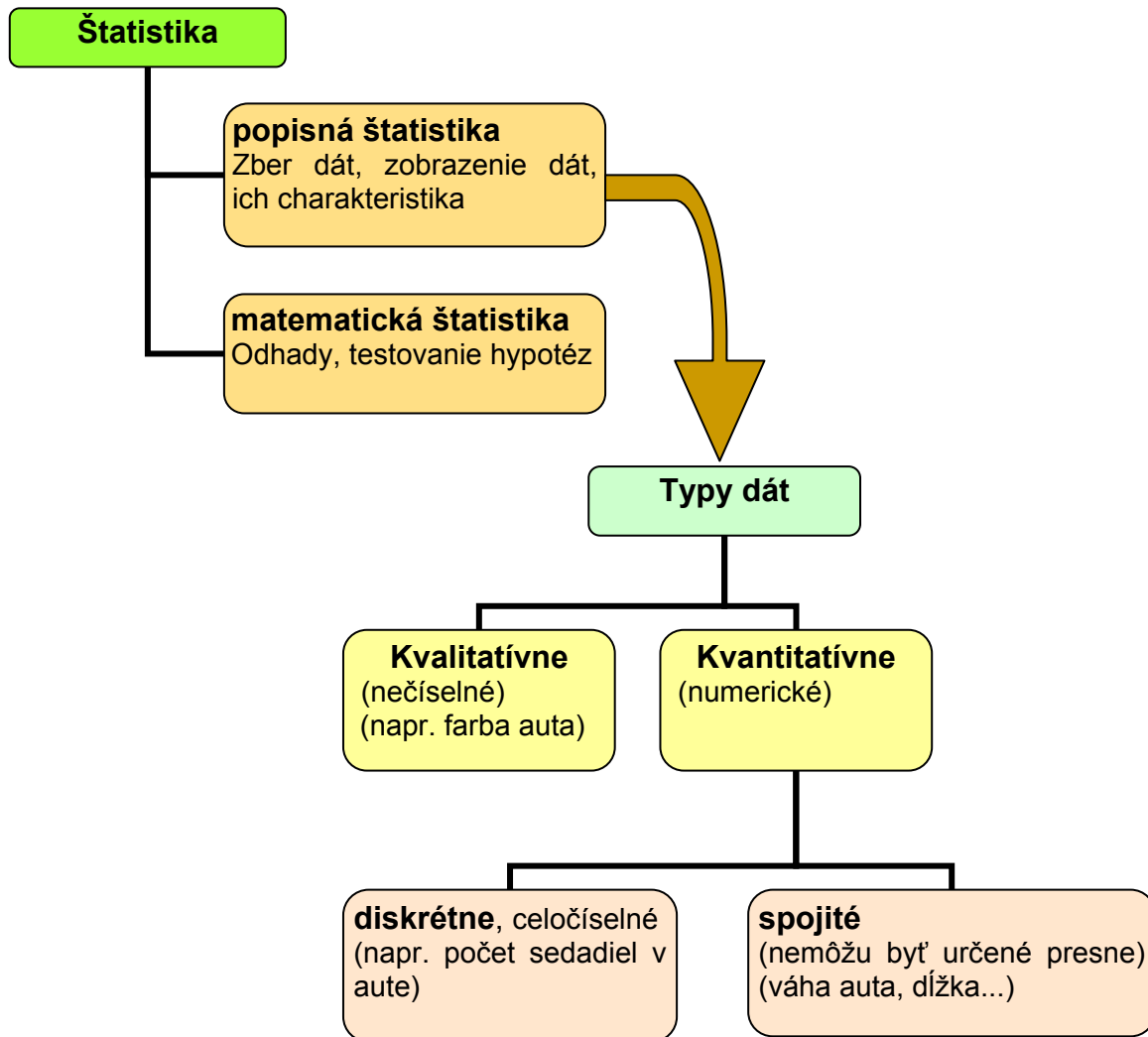
Dáta, priemer, rozptyl, štandardná odchýlka, modus, medián, šikmosť, špicatosť, kovariancia, korelácia.

1.1 POPIS DÁT

Všetko sa odvíja od dát s ktorými chceme pracovať. Môžeme ich nazývať súborom výsledkov, vzorkami, eventami alebo inak, ale pozostávajú zo súboru základných meraní z ktorých chceme získať nejakú zmysluplnú informáciu. Aby sme názornejšie videli význam dát, potrebujeme si ich zobraziť, alebo previesť do jedného, dvoch dôležitých čísel, ktoré daný súbor dát charakterizujú bez hlbšej analýzy alebo interferencie. Týmto sa zaoberá popisná štatistika

1.1.1 Typy dát

Dáta sa nazývajú kvantitatívne, alebo numerické ak môžu byť zapísané číslami a kvalitatívne, alebo nečíselné, ak sa číslami nedajú zapísať (napr. farba auta). S kvalitatívnymi dátami sa pre ťažkosti s ich matematickým popisom nebudeme v ďalšom zaoberať. Kvantitatívne dáta môžu byť dvoch typov. Diskrétne, z ich



povahy vyplýva, že sa dajú vyjadriť celými číslami a spojité, ktorých hodnota sa vyjadruje reálnymi číslami. Spojité dáta nemôžu byť zaznamenané presne, je možné ich zaznamenať len na istý počet desatinných miest. Príkladom diskretných dát môže byť počet kolies, sedadiel v aute, zatiaľ čo dĺžka a váha auta sú spojité dáta.

Pokiaľ máme veľký počet dát, často je výhodné rozdeliť ich do binov (skupín, tried, alebo blokov) zostavením frekvenčnej tabuľky.

Príklad: Hádzanie 20 mincí

Hádzeme 20 mincí. Vyjadrite skráteno dosiahnuté výsledky.

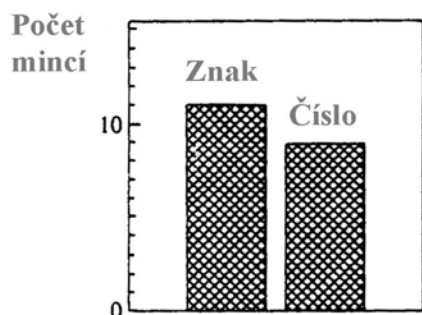
Riešenie

– každá minca môže po dopade ukazovať buď znak (Z) alebo číslo (C). Dostaneme tak napr. CCZCZZZZCCZCZCCZCZZZ, čo môžeme skráteno zapísať ako (11Z, 9C).

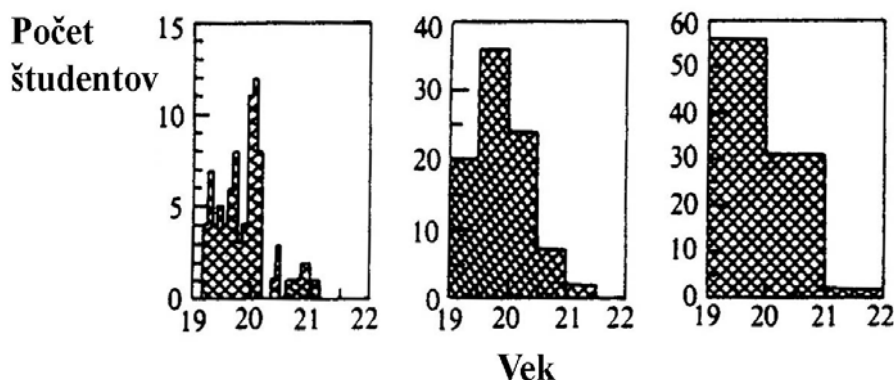
Pre spojité dáta je to komplikovanejšie, pretože všetky dáta sú rôzne (ak použijeme dostatočný počet desatinných miest). Ak chceme vytvárať skupiny čísel, musíme zvoliť určitý rozsah hodnôt, ktoré budú patriť do jedného binu. To znamená určité zaokrúhľovanie čísel a stratu presnosti informácie, čo je cena za kompresiu dát. Obvykle sa šírka binov vyberá rovnaká, ale niekedy môže byť výhodné použiť rozdielnu šírku binov.

1.1.2 Zobrazovanie dát

Počet eventov v každom bine môže byť zobrazený v stĺpcovom grafe alebo v histograme (obr. 1. 1). Medzi stĺpcovým grafom a histogramom je technický rozdiel v tom, že v prvom prípade je reprezentované číslo úmerné dĺžke stĺpca, zatiaľ čo v druhom prípade ploche. Stĺpcový graf je možné použiť pre kvalitatívne ako aj kvantitatívne dáta, zatiaľ čo histogram len pre kvantitatívne dáta. V prípade numerických dát je dôležitý správny výber šírky binu aby sme získali optimálne zobrazenie požadovaných charakteristík dát. Ak je šírka príliš malá, počet eventov v bine je malý a výraznejšie sa prejavujú fluktuácie dát, ak je šírka binu príliš veľká, stráca sa informácia o rozdelení dát. Dokumentuje to aj príklad (obr.1.2) rozdelenia skupiny študentov podľa veku.



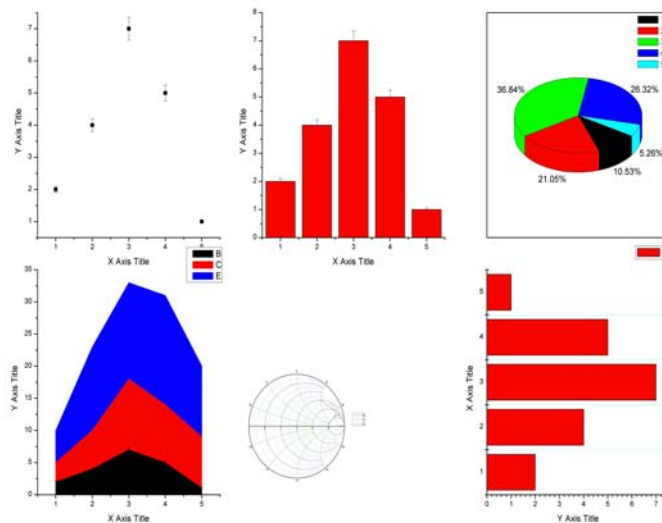
Obr.1.1 Príklad stĺpcového grafu zobrazujúceho výsledky hodu 20 mincami



Obr. 1. 2 Vek skupiny študentov zobrazený s rôznou šírkou binu

Dáta je možné zobraziť mnohými spôsobmi. Je veľa typov grafov (obr. 1.3), (napr. kruhový, bodový, čiarový graf, frekvenčný polygónový a veľa ďalších), takže je

potrebné vybrať také zobrazenie, ktoré optimálne zdôrazí požadovanú charakteristiku súboru dát.



Obr.1. 3 Príklad rôznych typov zobrazenia dát

1.2 PRIEMERY

1.2.1 Aritmetický priemer

Ak chceme popísať dáta jedným číslom, najlepším a najzmyslupnejším sa javí aritmetický priemer. Ak máme súbor dát obsahujúci N hodnôt $x \{x_1, x_2, \dots, x_N\}$, potom priemerná hodnota \bar{x} je

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

Takýmto spôsobom môžeme vypočítať aj strednú hodnotu ľubovolnej funkcie $f(x)$

$$\bar{f} = \frac{1}{N} \sum_{i=1}^N f(x_i)$$

Ak máme binované (zskupené) dáta, bin j zodpovedá hodnote x_j a obsahuje n_j dát, potom tieto priemery môžeme písať ako

$$\bar{x} = \frac{1}{N} \sum_{j=1}^N n_j x_j \quad , \quad \text{pre funkcie} \quad \bar{f} = \frac{1}{N} \sum_{j=1}^N n_j f(x_j)$$

Príklad: Váženie kovových platničiek

Hmotnosti 4 kovových platničiek sú 35g, 21g, 29g, 27g. Aká je priemerná hmotnosť platničky?

Riešenie

Priemer vypočítame podľa vzťahu $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$. Po dosadení dostaneme $(35+21+29+27) / 4 = 28$ g.

Priemerná hmotnosť potom bude 28 g.

Príklad: Obsadenosť áut

Na diaľnici sledujeme počet pasažierov v 100 autách. V 68 je len jeden pasažier, v 20 dvaja, v 8 traja a 4 štyria. Aká je priemerná obsadenosť auta pasažiermi?

Riešenie

Vychádzame zo vzťahu $\bar{x} = \frac{1}{N} \sum_{j=1}^N n_j x_j$. Po dosadení dostaneme $(68 \times 1 + 20 \times 2 + 8 \times 3 + 4 \times 4) / 100 =$

1.48. Priemerná obsadenosť auta potom bude 1.48 pasažiera.

1.2.2 Alternatívy aritmetickému priemeru

Geometrický priemer dvoch čísel je dĺžka strany štvorca s plochou rovnou súčinu dvoch čísel. Pre N čísel je potom definovaný ako

$$\sqrt[N]{x_1 x_2 x_3 \dots x_N}$$

Harmonický priemer je prevrátená hodnota aritmetického priemeru prevrátených hodnôt

$$\frac{N}{1/x_1 + 1/x_2 + \dots + 1/x_N}$$

Kvadratický priemer (odmocnina priemeru kvadrátov (root mean square)) môžeme vyjadriť ako

$$\sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_N^2}{N}}$$

Všetky tieto veličiny sú oveľa menej bežné ako aritmetický priemer, takže obvykle, keď sa hovorí o priemere, myslí sa tým aritmetický priemer.

Súbor dát sa často charakterizuje aj pomocou modusu a mediánu

Modus

je najčastejšie vyskytujúca sa hodnota v súbore dát. Dá sa ľahko nájsť, ale nemusí veľmi charakterizovať daný súbor dát, preto netreba preceňovať jeho význam. Súbor môže mať aj viac modulusov, alebo žiadny.

Medián

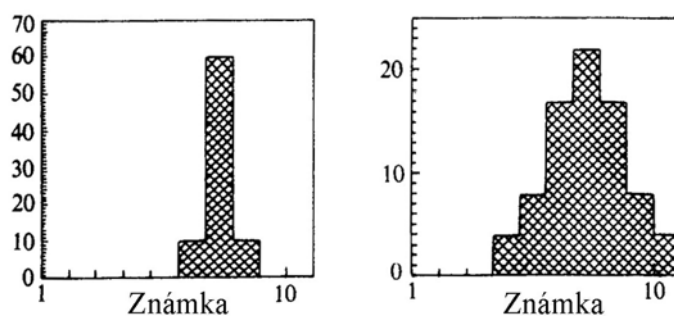
je bod v polovici súbore dát, rozpoluje súbor dát, polovica dát je nad mediánom a polovica je pod mediánom. Pokiaľ je dôležitejší rozsah dát ako ich číselné hodnoty, býva medián preferovaný pred aritmetickým priemerom. Medián dostaneme ako číselnú hodnotu v polovici podľa veľkosti usporiadaného súbore dát. Pokiaľ počet dát

je párnny, mediánom je priemerná hodnota dvoch dát v polovici usporiadaného súboru dát.

1.3 MERANIE ROZMAZANIA

1.3.1 Rozptyl

Priemer (stredná hodnota) charakterizuje naše dáta jedným číslom. Toto však môže byť niekedy nedostatočné a zavádzajúce, ako to dokumentuje príklad na obr. 1. 4. Oba dátové súbory vychádzajúce z hodnotenia 80 študentov 10 bodovou stupnicou majú rovnakú priemernú hodnotu (7), ale na prvý pohľad sa výrazne líšia.



Obr.1.4 Hodnotenie skupiny 80 študentov dvomi učiteľmi bodmi z 10 stupňovej škály

Prvý učiteľ hodnotil vo veľmi malom rozpätí, zatiaľ čo druhý použil podstatne širšiu škálu hodnotenia. Sú rozdelené rozdielne a je potrebné nejakým spôsobom charakterizovať ich šírku, alebo rozmazanie dát okolo strednej hodnoty. Použitie priemernej odchýlky nie je vhodné, pretože kladné a záporné odchýlky sa rušia a výsledok je nulový.

$$\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) = \frac{1}{N} \sum_{i=1}^N x_i - \frac{1}{N} \sum_{i=1}^N \bar{x} = \bar{x} - \bar{x} = 0$$

Výhodnejšie je použiť priemer kvadrátov odchýliek – rozptyl (variance)

$$V(x) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

Pre funkcie môžeme rozptyl vyjadriť ako

$$V(f) = \frac{1}{N} \sum_{i=1}^N (f(x_i) - \bar{f})^2$$

Definícia rozptylu môže byť jednoduchými operáciami prevedená do jednoduchšieho tvaru

$$V(x) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

$$V(x) = \frac{1}{N} \sum_{i=1}^N (x_i^2 - 2x_i\bar{x} + \bar{x}^2)$$

$$V(x) = \frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{1}{N} \sum_{i=1}^N 2x_i\bar{x} + \frac{1}{N} \sum_{i=1}^N \bar{x}^2$$

$$V(x) = \frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{1}{N} 2\bar{x} \sum_{i=1}^N x_i + \frac{1}{N} \bar{x}^2 \sum_{i=1}^N 1$$

$$V(x) = \overline{x^2} - 2\bar{x}^2 + \bar{x}^2$$

Takže nakoniec dostaneme

$$V(x) = \overline{x^2} - \bar{x}^2$$

alebo ekvivalentne

$$V(x) = \frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2$$

Slovne to môžeme vyjadriť, že rozptyl je priemer kvadrátov mínus kvadrát priemerov.

1.3.2 Štandardná odchýlka

Stredná kvadratická odchýlka sa nazýva štandardná odchýlka a označuje sa σ . Je definovaná ako odmocnina z rozptylu

$$\sigma = \sqrt{V(x)}$$

$$\sigma = \sqrt{\overline{x^2} - \bar{x}^2}$$

alebo inak vyjadrené

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2}$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Príklad: Hodnotenie študentov

Porovnajcie hodnotenia tej istej skupiny študentov dvomi učiteľmi na obr. 1.4. Čím sa líšia?

Riešenie

Na obr.1.4 sú ukázané dve rôzne hodnotenia tých istých študentov dvomi učiteľmi - prvý učiteľ udelil 10 študentom hodnotenie 6, 60-tim hodnotenie 7, 10 - 8, pričom priemerné hodnotenie bolo 7, rozptyl 0.25, štandardná odchýlka 0.50. Druhé hodnotenie bolo jemnejšie - učiteľ použil väčší rozsah známok, pričom priemerné hodnotenie bolo tiež 7, ale rozptyl bol 2.1 dosiahnutá štandardná odchýlka bola 1.45 (čo je ~3x viac ako v prvom prípade)

Príklad: Použitie štandardnej odchýlky pri kontrole kvality produkcie

Spoločnosť vyrába ložiská s priemernou hmotnosťou 30 g so štandardnou odchýlkou 0.1g. Ako je možné využiť štandardnú odchýlku pri výstupnej kontrole?

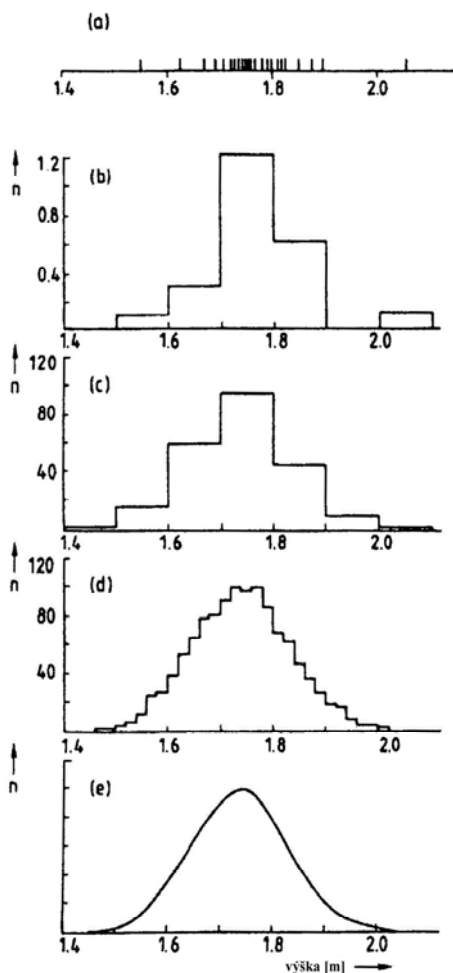
Riešenie

Pokiaľ výrobky majú hmotnosti v rámci jednej štandardnej odchýlky je to v poriadku, výrobný proces beží bez problémov. Výstupná kontrola ešte akceptuje ložiská s váhou 29.8 až 30.2 g (2σ) ale je už

potrebné sledovať podrobnejšie výrobný proces - dve štandardné odchýlky reprezentujú výstražnú hladinu, keď váha presahuje 3σ (rozsah 29.7 a 30.3g) je to už hladina akčná – produkcia je zastavená

1.3.2.1 Iné definície štandardnej odchýlky

Obr.1. 5 ukazuje príklad možného rozdelenia spojitej premennej z niekoľkých pozorovaní napr. výšky 30-ročných mužov. Ak je dostupných len niekoľko hodnôt, dáta môžeme zobrazit' čiarkami na osi výšky (obr.1.5a), alebo zobrazit' ako histogram, kde na zvislej osi je počet pozorovaní v danom intervale výšok Δh , na vodorovnej osi výšok má stĺpec šírku 10 cm (obr. 1.5b). Aktuálny počet mužov v danom bine je $n\Delta h$, celkový počet je $\sum n\Delta h$. Ak zväčšíme počet pozorovaní 100 krát, počet v každom bine sa zvýši (obr. 1.5c) a je možné vykreslit' histogram s menšou šírkou binu aby sme zobrazili tvar rozdelenia s lepším rozlíšením. Pretože zobrazujeme $n(h)$ ako počet pozorovaní na v závislosti na výške v cm, bez ohľadu na šírku binu, celková výška histogramu sa nebude veľmi menit' pri zmene šírka binu (obr. 4d). Pre veľký počet pozorovaní môže byť šírka binu veľmi malá, takže



Obr.1.5 Príklad možného rozdelenia spojitej premennej (napr. výšky 30-ročných mužov). - (a) len z niekoľkých pozorovaní – každé reprezentuje čiarka na osi výšky. Tieto dáta môžu byť alternatívne znázornené histogramom (b), kde n je počet mužov v danom intervale výšok (šírka binu je tu 10 cm). (c) – počet pozorovaní 100 krát vyšší ako v predošlom prípade. (d) – ten istý počet pozorovaní ako v (c), ale s menšou šírkou binu. (e) – so zvyšovaním počtu pozorovaní a znižovaním šírky binu prejdeme k spojitému rozdeleniu. (Zdroj: L. Lyons, Statistics for nuclear and particle physics, Cambridge University Press, 1986)

histogram môžeme aproximovať spojitou krivkou (obr. 1.5e). Takýto prípad nám bude dávať veľmi dobrú teoretickú predpoveď počtu mužov rôznych výšok, pričom celkový počet mužov zahrnutých v histograme bude $\int n(h)dh$.

Príjmime konvenciu, že skutočnú strednú hodnotu budeme označovať μ a strednú kvadratickú odchýlku σ . Niektorí autori definujú štandardnú odchýlku ako r.m.s. (root

mean square) - odchýlku dát od skutočnej strednej hodnoty μ miesto od vzorkového priemeru \bar{x}

$$\mu = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_i x_i \right) = \langle x \rangle \quad \sigma = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \quad \sigma = \sqrt{\frac{1}{N} \sum_i (x_i - \mu)^2}$$

Merané hodnoty priemeru a rozptylu určitej (obmedzenej) vzorky budeme označovať \bar{x} a s^2 . μ obvykle nevieme, vieme len jeho odhad \bar{x} (priemer vzorkovej populácie). Aby sme kompenzovali nadhodenie, znížime počet stupňov voľnosti. Ak nahradíme $x - \mu$, budeme robiť najmenej jedno meranie na určenie x . Najlepší experimentálny odhad štandardnej odchýlky σ bude vzorková štandardná odchýlka s . Nevychýlený (unbiased) odhad σ^2 zdrojového (skutočného) rozdelenia s je daný ako

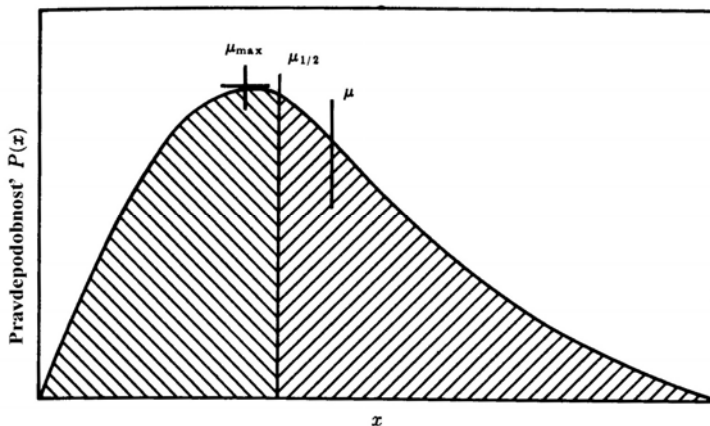
$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

alebo po jednoduchých úpravách dostaneme ekvivalentne

$$s^2 = \frac{N}{N-1} (\overline{x^2} - \bar{x}^2) \quad , \quad \text{kde} \quad \overline{x^2} = \frac{1}{N} \sum_i x_i^2 \quad \text{a} \quad \bar{x}^2 = \left(\sum_i x_i / N \right)^2$$

Keď sa bude počet meraní zvyšovať, nakoniec pre nekonečný počet meraní dosiahneme pôvodné, zdrojové (parent) rozdelenie meranej veličiny

zdrojový parameter = $\lim_{N \rightarrow \infty}$ (experimentálny parameter)



Obr.1.6 Príklad typickej funkcie rozdelenia pravdepodobnosti ukazujúci najpravdepodobnejšiu hodnotu μ_{\max} , medián $\mu_{1/2}$ a strednú hodnotu μ . Dve rôzne vyšrafované časti sú rovnaké

Faktor $1/(N-1)$ je potrebný na to, aby sme dostali nevychýlený odhad populačného rozptylu σ^2 , t.j. pre veľké vzorky s^2 sa blíži k σ^2 .

Ak merania sú zoskupené, t.j. ak hodnotu x_j má m_j eventov, dostaneme

$$\bar{x} = \sum_j m_j x_j / \sum_j m_j \quad \text{a}$$

$$s^2 = \sum_j m_j (x_j - \bar{x})^2 / (\sum_j m_j - 1) .$$

Rozptyl priemeru je

$$u^2 = s^2 / \sum_j m_j$$

Príklad: Meranie s rozdielnou detekčnou účinnosťou
 Predpokladajme, že individuálne merania boli robené s rozdielnou detekčnou účinnosťou (buď kvôli meraniu s rôznymi časťami prístroja, alebo kvôli tomu, že vlastná účinnosť je funkciou premennej x). Ako je potrebné tento fakt zobrať do úvahy?

Riešenie

Každé meranie x_k je nutné váhovať faktorom w_k ktorý je recipročný detekčnej účinnosti pre dané meranie. Potom priemer je daný ako:

$\bar{x} = \frac{\sum_k w_k x_k}{\sum_k w_k}$. Pre rozptyl a rozptyl priemeru (u^2) dostaneme:

$$s^2 = \left(\frac{\sum_k w_k (x_k - \bar{x})^2}{\sum_k w_k} \right) \times n_{\text{eff}} / (n_{\text{eff}} - 1) \quad \text{a} \quad u^2 = s^2 / n_{\text{eff}}.$$

n_{eff} tu je efektívny počet eventov. Ak celkový počet eventov je $T \pm \delta$, potom

$$n_{\text{eff}} = \frac{T^2}{\delta^2} = \left(\frac{\sum_k w_k}{\sum_k w_k^2} \right)^2. (**)$$

Teda 100 ± 10 eventov nám dá 100 efektívnych eventov (zo vzťahu (**)). Ale ak máme konštantnú detekčnú účinnosť 4%, registrujeme len 4 ± 2 reálnych eventov, korigovaný celkový počet je 100 ± 50 a n_{eff} je 4 (ako očakávame). Na záver, dá sa očakávať, že vzorky eventov, ktoré majú veľmi nízku detekčnú účinnosť, budú dávať pre n_{eff} hodnotu ~ 1 .

1.3.3 Alternatívne merania rozmazania

stredná absolútna odchýlka - je veľmi zriedkavo používaná

$$\frac{1}{N} \sum_i |x_i - \bar{x}|$$

rozsah dát (range) - rozdiel medzi najvyššou a najnižšou hodnotou (variačné rozpätie)

interkvartilový rozsah (interquartile range) – (kvartilové rozpätie) – rozdiel medzi spodným a horným kvartilom

Kvartily (quartile) – dáta (usporiadané) rozdelené na 4 časti

spodný kvartil - bod s 25 % dát pod túto hodnotu a 75% dát nad

horný kvartil - bod s 25 % dát nad a 75% dát pod

Niekedy sa tiež používajú tiež decily obsahujúce 10% dát, alebo percentily, ktoré obsahujú dané percento dát. So štandardnou odchýlkou σ niekedy bývajú problémy - jej hodnota môže byť majorizovaná niekoľkými extrémnymi hodnotami na okrajoch (chvostoch) dát, definuje sa FWHM (plná šírka v polovičnej výške) – táto nezávisí na dátach na chvostoch, používa len centrálny pík (dá sa veľmi ľahko nájsť z histogramu použitím ceruzky a pravítka). Pre Gaussove rozdelenie platí nasledovný vzťah medzi FWHM a strednou kvadratickou odchýlkou

$$\text{FWHM} = 2.35\sigma$$

1.4 VLASTNOSTI KVANTITATÍVNYCH DÁT

1.4.1 Miery polohy (central tendency - location)

(priemer, modus, medián)

1.4.2 Miery variability (variation - dispersion)

(variačné rozpätie, kvartilové rozpätie, variačný koeficient, štandardná odchýlka, rozptyl)

1.4.3 Miery tvaru (shape) (šikmosť, špicatosť);

1.4.3.1 Šikmosť (skew)

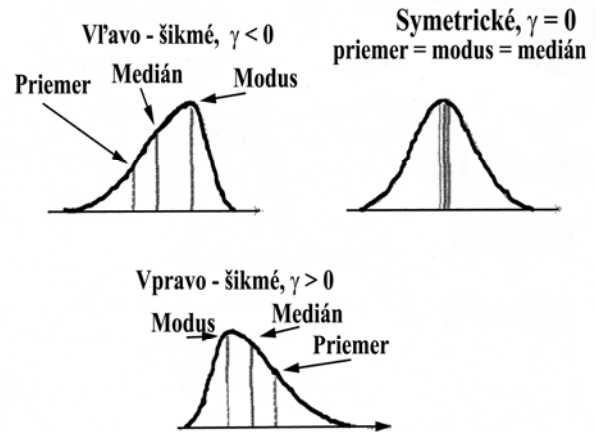
popisuje asymetriu dát

$$\gamma = \frac{1}{N\sigma^3} \sum_i (x_i - \bar{x})^3 = \frac{1}{\sigma^3} \overline{(x_i - \bar{x})^3}$$

po úpravách dostaneme

$$\gamma = \frac{1}{\sigma^3} (\overline{x^3} - 3\bar{x}\overline{x^2} + 2\bar{x}^3)$$

γ je nulové ak dáta sú symetrické, rozdelené okolo priemeru, ak chvost je roztiahnutý napravo γ je pozitívne, dáta majú pozitívnu šikmosť, negatívnu majú ak deformácia je naľavo. Maxwellove rozdelenie rýchlostí molekúl v plyne - príklad šikmosť rozdelenia. Alternatívy definovania šikmosť - Pearsonova šikmosť (Pearson skew)

$$\text{šikmosť} = \frac{\text{priemer} - \text{modus}}{\sigma}$$


Obr.1.7 Ilustrácia šikmosťi

1.4.3.2 Špicatosť (curtosis, niekedy tiež kurtosis)

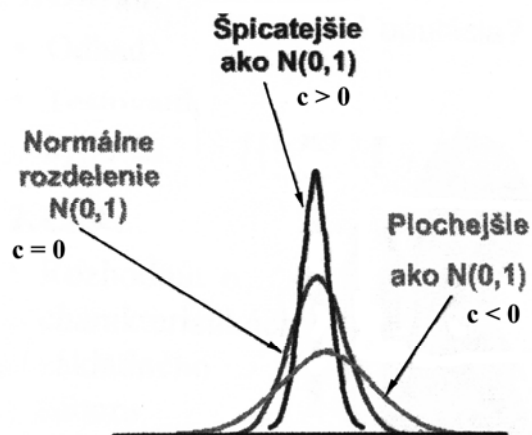
$$c = \frac{1}{\sigma^4} \overline{(x - \bar{x})^4} - 3$$

pre normálne rozdelenie c je rovné nule. Pozitívne c implikuje relatívne vyššie píky širšie chvosty ako Gauss s rovnakým priemerom a štandardnou odchýlkou. Záporné hodnoty znamenajú širší pík a kratšie chvosty

Vo všeobecnosti:

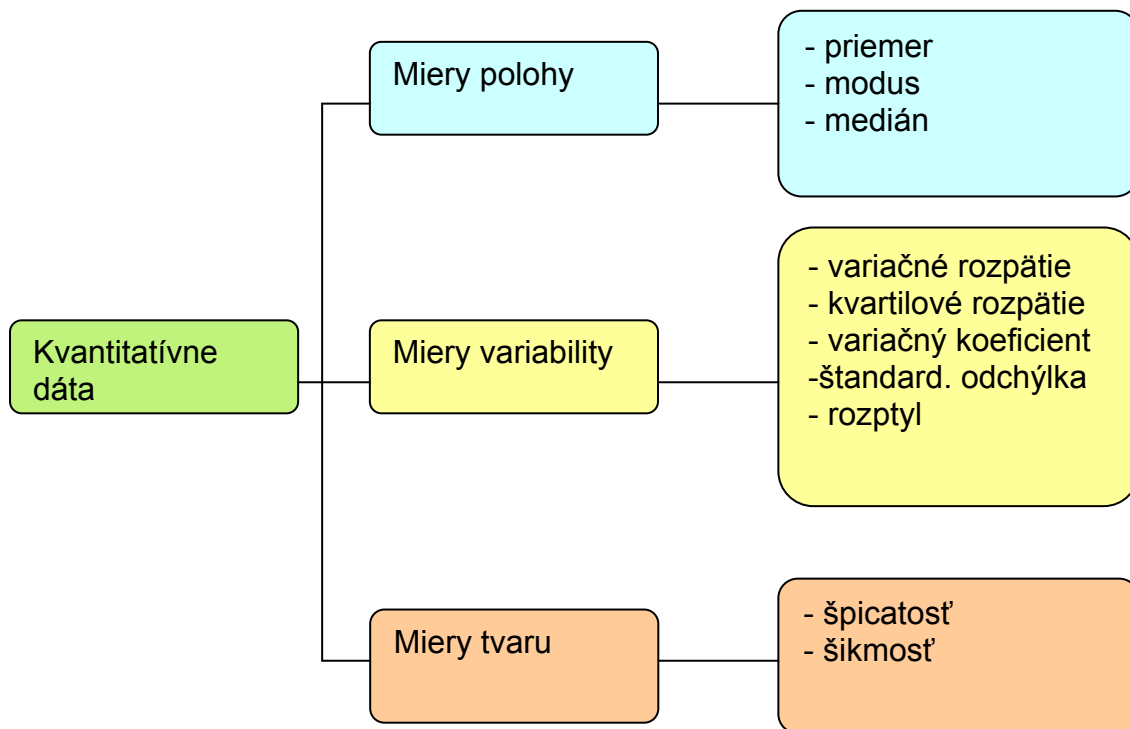
$$\frac{1}{N} \sum_{i=1}^N x_i^r \quad \text{je } r\text{-tý moment } x \text{ a}$$

$\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^r$ sa volá r -tý **centrálny** moment x .



Obr. 1.8 Ilustrácia špicatosťi súboru dát

Rozptyl je druhý, šikmosť tretí a curtosis je štvrtý centrálny moment



1.5 KOVARIANCIA A KORELÁCIA

Vyjadrujú vzťah medzi dvojicami dát v súbore dát, napr. dvojice polohu a čas (x,t), napr. výška a váha, IQ atď pre triedu študentov, pre každého jednotlivca a pod.

1.5.1 Kovariancia

Predpokladajme dáta pozostávajúce z párov čísel $\{(x_1, y_1), (x_2, y_2), \dots\}$, Môžeme nájsť ich priemer, rozptyl $V(x)$ a $V(y)$, a štandardné odchýlky σ_x, σ_y

Dáta môžu spolu súvisieť – môžu byť vzájomne nezávislé alebo závisieť na sebe. Popisuje to kovariancia medzi x a y

$$\begin{aligned} \text{cov}(x, y) &= \frac{1}{N} \sum_i (x_i - \bar{x})(y_i - \bar{y}) \\ &= \overline{(x_i - \bar{x})(y_i - \bar{y})} \\ &= \overline{xy} - \bar{x} \bar{y} \end{aligned}$$

Ak hodnoty x, ktoré sú nadpriemerné, majú tendenciu vyskytovať sa spolu s nadpriemernými hodnotami y (čo implikuje, že malé x podobne bude sprevádzané malým y) potom znamienko oboch častí v sume bude mať tendenciu byť rovnaké a celková suma bude kladná. Podobne ak veľké x má tendenciu byť s malým y, kovariancia je záporná.

Ak hodnoty sú nezávislé a nespojené, potom kladný rozdiel pre x má rovnakú šancu byť násobený kladným alebo záporným rozdielom pre y a celková suma je nulová.

Všimnime si, že ak kovariancia obsahuje len rozdiely medzi priermi x a y , nič sa nemení ak sa posunie počiatok, teda vo všeobecnosti $\text{cov}(x, x) = V(x)$

Príklad: vzťah medzi výškou a váhou skupiny osôb

Aký je asi vzťah medzi výškou a váhou skupiny osôb?

Riešenie

Kovariancia medzi výškou a váhou skupiny osôb je podľa všetkého pozitívna, pretože vyšší ľudia obvykle vážia viac. Medzi váhou a silou môže byť negatívna, pretože príliš tuční ľudia obvykle nebývajú vo veľmi dobrej kondícii, medzi výškou a IQ bude kovariancia pravdepodobne nulová.

1.5.2 Korelácia

Kovariancia je užitočná, ale má rozmer (napr. kovariancia medzi výškou a váhou 6.5 má jeden význam je ak sú rozmery [cm] - [g], iný význam ak sú [m] - [kg]).

Lepšie je vyjadriť vzťahy medzi premennými bezrozmerným korelačným koeficientom, ktorý je definovaný ako

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{xy} - \bar{x} \bar{y}}{\sigma_x \sigma_y}$$

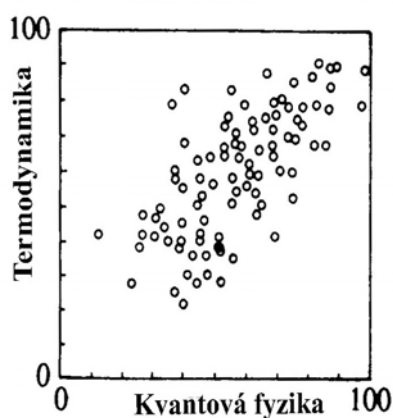
kde ρ je číslo medzi -1 a +1, ak je rovný nule, x a y sú nekorelované. Kladná korelácia znamená, že ak dané x je väčšie ako priemerné x , potom aj y bude tiež (v priemere) väčšie ako priemerné y . Pre záporné ρ - väčšie x implikuje menšie y . Ak $\rho = 1$ (alebo -1) potom x a y sú úplne korelované (ak vieme hodnotu jedného, presne špecifikujeme hodnotu druhého). Korelačný koeficient je bezrozmerný. Obr. 1.11 ukazuje rôzne korelácie dát

Príklady korelácie

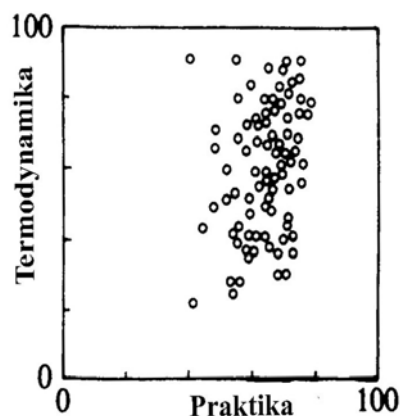
Na obr. 1.9 a 1.10 sú ukázané výsledky skúšky študentov druhého ročníka istej anglickej univerzity z termodynamiky a kvantovej mechaniky so zahrnutím hodnotenia laboratórnych prác. V akom sú vzťahu?

Riešenie

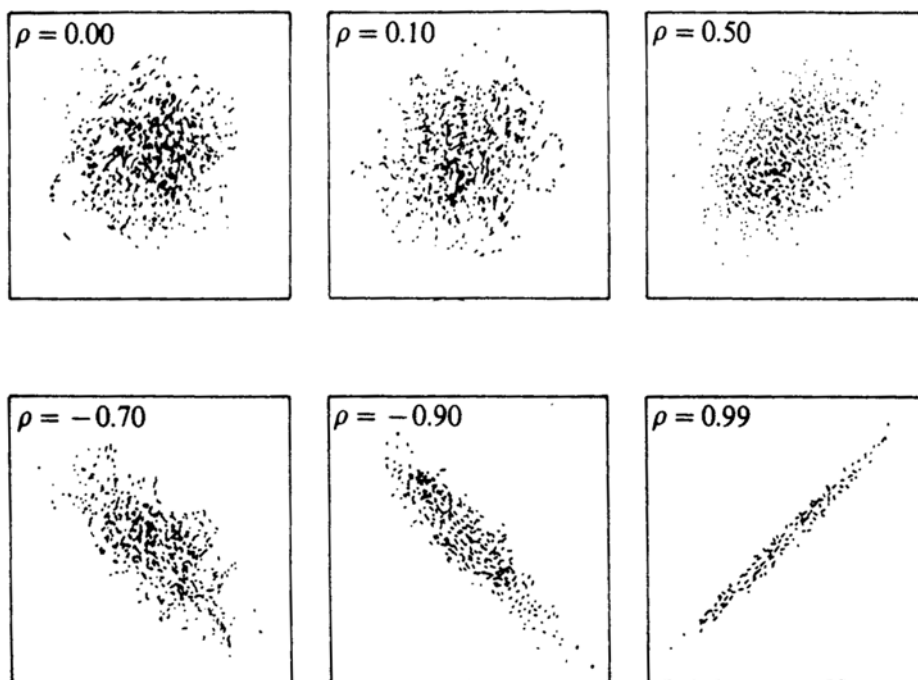
Obr. 1.9 zobrazuje známky z kvantovej mechaniky a termodynamiky, vidieť, ako sa dá očakávať, že kto je dobrý v jednom je dobrý aj v druhom - korelačný koeficient 0.7. Obr. 1.10 ukazuje porovnanie výsledkov termodynamiky vs. praktické cvičenia. Tendencia kto je dobrý v jednom je dobrý aj v druhom je prítomná, ale je veľmi slabá - korelačný koeficient 0.3



Obr. 1.9 Porovnanie hodnotenia študentov z termodynamiky a kvantovej fyziky



Obr. 1.10 Porovnanie hodnotenia študentov z termodynamiky a praktických cvičení



Obr.1.11 Príklady korelácie dát s rôznymi hodnotami korelačného koeficientu

1.5.3 Viac než dve premenné

Ak máme tri elementy - (x, y, z) označovanie je vcelku bez problémov, ale pre n elementov už to nie je také jednoduché. Jedným z možných riešení je označovanie $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$.

Kovariancia sa vyjadruje medzi párom premenných

$$\text{cov}(x_{(i)}, y_{(j)}) = \overline{x_{(i)}y_{(j)}} - \bar{x}_{(i)}\bar{y}_{(j)}$$

S týmto označovaním vytvoria tieto elementy symetrickú maticu n x n prvkov

$$V_{ij} = \text{cov}(x_{(i)}, x_{(j)})$$

táto sa nazýva kovariančná matica, alebo rozptylová matica, alebo chybová matica, a diagonálne prvky sú práve rozptyly. Korelačná matica je bezrozmerný ekvivalent kovariančnej matice. Jej prvky musia ležať medzi -1 a +1 a ukazujú mieru korelovanosti dvoch premenných. Jej diagonálne prvky sú všetky = 1

$$\rho_{ij} = \frac{\text{cov}(x_{(i)}, x_{(j)})}{\sigma_i \sigma_j}$$

kovariančnú maticu môžeme zapísať v tvare

$$V_{ij} = \rho_{ij} \sigma_i \sigma_j$$

KONTROLNÉ OTÁZKY:

1. Aké je rozdelenie štatistiky a čím sa zaoberajú jej jednotlivé časti?
2. Aké typy dát poznáme a ako ich môžeme charakterizovať?
3. Ako môžeme dáta zobrazovať?
4. Čím môžeme vo väčšine prípadov najvhodnejšie charakterizovať súbor dát pokiaľ to máme vyjadriť jedným číslom?
5. Ako vypočítame aritmetický priemer pre jednotlivé dáta a ako pre zoskupené (binované) dáta?
6. Čo je to modus a medián a ako ich vypočítame?
7. Čo je to rozptyl a ako ho vypočítame?
8. Ako vypočítame štandardnú odchýlku?
9. Čo vyjadruje vzorková štandardná odchýlka a ako ju vypočítame?
10. Čo sú kvartily a čo je interkvartilové rozpätie?
11. Aký je vzťah medzi FWHM a štandardnou odchýlkou pre Gaussove rozdelenie?
12. Ako môžeme charakterizovať vlastnosti kvantitatívnych dát?
13. Čo vyjadruje šikmosť (skew) a ako ju vypočítame?
14. Čo je to špicatosť (kurtosis) a ako ju vypočítame?
15. Čo je to kovariancia a ako ju vypočítame?
16. Čo vyjadruje korelácia a ako ju vypočítame?
17. Aký je rozdiel medzi kovarianciou a koreláciou?

ÚLOHY:

- 1.1 Vek (v rokoch) triedy 23 študentov je nasledovný: 18.9, 18.6, 19.4, 18.8, 19.3, 20.3, 19.8, 20, 18.5, 19.5, 19.1, 18.6, 19.4, 19.2, 18.8, 18.6, 18.2, 19.8, 19.5, 19, 20, 19.6, 18.5. Vypočítajte priemer a štandardnú odchýlku.
- 1.2 Zopakujte výpočet pre príklad 1. so zahrnutím veku učiteľa (37.0). Všimnite si, že efekt na priemer je veľmi slabý, ale štandardná odchýlka vzrástla významne.
- 1.3 Vypočítajte pre predošlé problémy šikmosť
- 1.4 Hodnotenia 12 študentov z klasickej mechaniky a kvantovej mechaniky sú: Klasická : 22 48 76 10 22 4 68 44 10 76 14 56; Kvantová: 63 39 61 30 51 44 74 78 55 58 41 69. Vypočítajte dve priemerné známky, kovarianciu a koreláciu.
- 1.5 Máme 80 čísel:
90 90 79 84 78 91 88 90 85 80
88 75 73 79 78 79 67 83 68 60
73 79 69 74 76 68 72 72 75 60
61 66 66 54 71 67 75 49 51 57

62 64 68 58 56 79 63 68 64 51
 58 53 65 57 59 65 48 54 55 40
 49 42 36 46 40 37 53 48 44 43
 35 39 30 41 41 22 28 36 39 51

Histogramujte tieto čísla použitím vhodného rozmeru binu.

1.6 Najdite priemer, modus, a medián pre uvedené dáta, jednak pre pôvodné ako aj pre binované.

1.7 Použitím predošlých dát vypočítajte štandardnú odchýlku a FWHM.

SÚHRN

Dáta - kvantitatívne, alebo numerické ak môžu byť zapísané číslami a kvalitatívne, alebo nečíselné, ak sa číslami nedajú zapísať (napr. farba auta). Kvantitatívne dáta môžu byť diskrétné, z ich povahy vyplýva, že sa dajú vyjadriť celými číslami a spojité, ich hodnota sa vyjadruje reálnymi číslami.

Aritmetický priemer $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$

Modus je najčastejšie vyskytujúca sa hodnota v súbore dát.

Medián je bod v polovici súboru dát, rozpoluje súbor dát. Medián dostaneme ako číselnú hodnotu v polovici podľa veľkosti usporiadaného súboru dát. Pokiaľ počet dát je párny, mediánom je priemerná hodnota dvoch dát v polovici usporiadaného súboru dát.

Rozptyl $V(x) = \overline{x^2} - \bar{x}^2$ alebo ekvivalentne $V(x) = \frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2$

Štandardná odchýlka $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$, Najlepší experimentálny odhad

štandardnej odchýlky σ bude **vzorková štandardná odchýlka s**. Nevychýlený odhad σ^2 zdrojového (skutočného) rozdelenia s je daný ako $s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$

Rozsah dát (range) - rozdiel medzi najvyššou a najnižšou hodnotou (varičné rozpätie)

Interkvartilový rozsah (interquartile range) – (kvartilové rozpätie) – rozdiel medzi spodným a horným kvartilom

Kvartily (quartile) – dáta (usporiadané) rozdelené na 4 časti, spodný kvartil - bod s 25 % dát pod túto hodnotu a 75% dát nad touto hodnotou, horný kvartil - bod s 25 % dát nad a 75% dát pod touto hodnotou

Šikmost' (skew) - popisuje asymetriu dát $\gamma = \frac{1}{N\sigma^3} \sum_i (x_i - \bar{x})^3 = \frac{1}{\sigma^3} \overline{(x_i - \bar{x})^3}$

Špicatosť (curtosis, niekedy tiež kurtosis) $c = \frac{1}{\sigma^4} \overline{(x - \bar{x})^4} - 3$, pre normálne rozdelenie c je rovné nule

Kovariancia medzi x a y $\text{cov}(x, y) = \frac{1}{N} \sum_i (x_i - \bar{x})(y_i - \bar{y}) = \overline{xy} - \bar{x}\bar{y}$

Korelácia - bezrozmerný korelačný koeficient - $\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}$, kde ρ je číslo medzi -1 a +1, ak je rovný nule , x a y sú nekorelované.

2 TEORETICKÉ ROZDELENIA

UČEBNÉ CIELE

Študent by sa mal oboznámiť s pojmom pravdepodobnosti, pôsobením zákona veľkých čísel, čo sú to očakávané hodnoty, zoznámiť sa so základnými rozdeleniami hustôt pravdepodobnosti pre binomické, Poissonove a Gaussove rozdelenie a ich charakteristikami, limitnými prípadmi vedúcimi ku Gaussovskému rozdeleniu.

KLÚČOVÉ SLOVÁ

Pravdepodobnosť, zákon veľkých čísel, očakávané hodnoty, binomické rozdelenie, Poissonove rozdelenie, Gaussove rozdelenie, binormálne rozdelenie

2.1 PRAVDEPODOBNOŠŤ

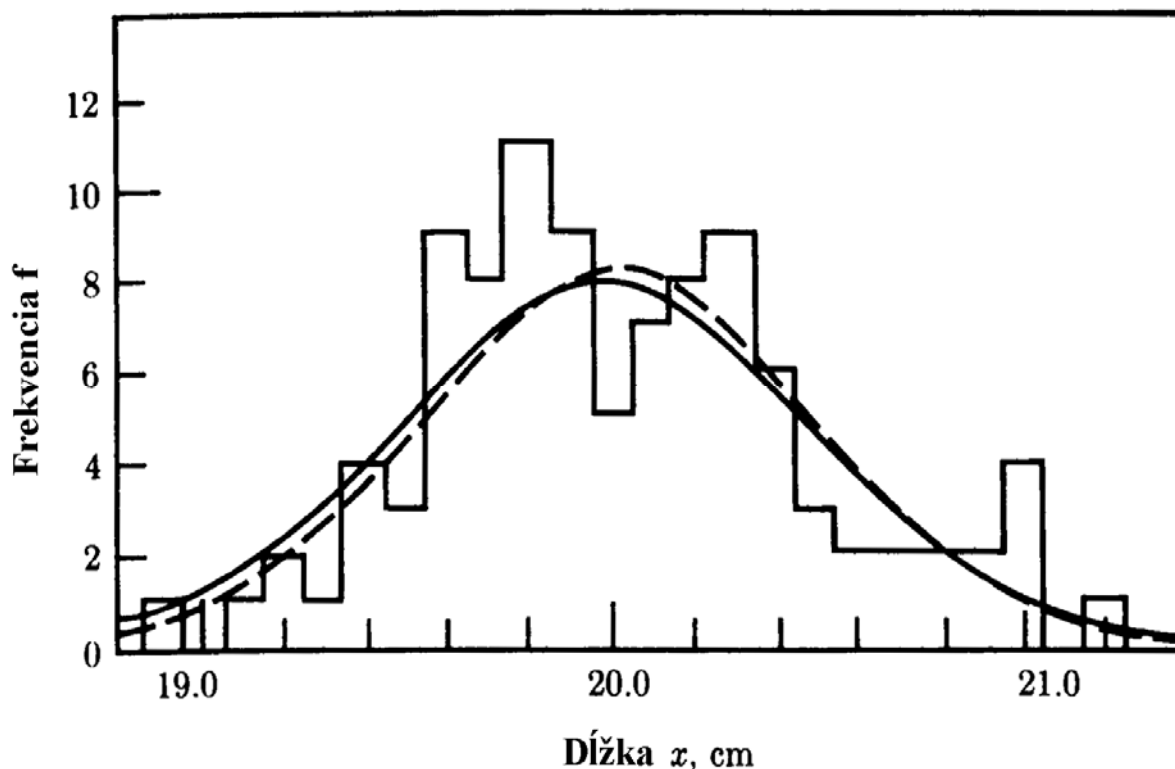
Čo je to pravdepodobnosť? V mnohých situáciách sa stretávame s experimentami v ktorých sa podstatné podmienky nemenia (zostávajú zachované) a napriek tomu v experimente dostávame rôzne výsledky. Takéto výsledky individuálnych meraní alebo experimentov môžu byť nepredpovedateľné, a predsa možné výsledky série takých meraní majú dobre definované rozdelenie. Pravdepodobnosť p získania určitých špecifických výsledkov realizáciou jedného z týchto meraní je možné vyjadriť pomerom

$$p = \frac{\text{počet udalostí v ktorých pozorujeme výsledok}}{\text{celkový počet meraní}}$$

S pojmom pravdepodobnosti sú spojené pojmy - možnosť, nemožnosť, istota. Pokiaľ pravdepodobnosť daného javu je menšia ako 1 ale väčšia ako 0 – jav je možný, pokiaľ pravdepodobnosť je rovná 0 – jav je nemožný a pokiaľ pravdepodobnosť je rovná 1 je istota, že jav určite nastane. Javy môžu byť vzájomne nezávislé na sebe, môžu sa vzájomne vylučovať alebo jeden jav môže byť podmienený existenciou iného javu. V prípade vzájomne nezávislých javov, pravdepodobnosť, že tieto javy

nastanú súčasne je daná súčinom pravdepodobností jednotlivých javov. Pokiaľ máme vzájomne sa vylučujúce javy, pravdepodobnosť, že nastane nejaký jav je daná súčtom pravdepodobností

vzájomne sa vylučujúce javy	- súčet pravdepodobností	$p = p_1 + p_2 + \dots + p_N$
nezávislé javy	- súčin pravdepodobností	$p = p_1 p_2 p_3 \dots p_N$
podmienené pravdepodobnosti	- $p = p_A \cdot p_{A B}$	



Obr. 2. 1 Graf závislosti frekvencie a meranej dĺžky pre 100 meraní. Plná čiara zobrazuje skutočné zdrojové rozdelenie so strednou hodnotou $\mu = 20.000$ cm a $\sigma = 0.50$ cm. Prerušovaná krivka zobrazuje odhad zdrojového rozdelenia s $\mu = 20.028$ cm a $\sigma = 0.48$ cm.

2.1.1 Zákon veľkých čísel

- spočíva v tom, že variabilita veľkého počtu nezávislých (alebo obmedzene závislých) náhodných veličín sa vzájomne natoľko ruší (kompenzuje), že ich súčet je relatívne (napr. vzhľadom k počtu sčítancov, alebo k svojej strednej hodnote), skoro konštantný. Inak povedané, ak rozmer súboru dát N vzrastá, fluktácie sa znižujú a náhodné výchylky sa vyhladzujú. Dokumentuje to aj obr. 2.1 kedy rozdelenie obmedzeného počtu meraní (100) je dosť rozhádzané, pri rastúcom počte meraní sa bude vyhladzovať.

Príklad platnosti zákona veľkých čísel je napr. priebeh chemických reakcií, aj keď je výsledkom náhodného chovania jednotlivých molekúl, je možné vzhľadom k veľkému množstvu a malej závislosti zúčastnených molekúl predvídať pomocou diferenciálnych rovníc.

2.2 OČAKÁVANÉ HODNOTY

Ak vieme rozdelenie pravdepodobnosti pre nejaké číslo r , vieme ľahko vypočítať priemernú hodnotu, ktorú môžeme očakávať. Očakávaná hodnota r - označuje sa $\langle r \rangle$, alebo niekedy $E(r)$ (expected value).

$$\langle r \rangle = \sum_r rP(r).$$

Každá funkcia r má tiež svoju očakávanú hodnotu $\langle f \rangle = \sum_r f(r)P(r)$

Pre súčet funkcií platí

$$\langle f + g \rangle = \sum_r (f + g)P(r) = \sum_r fP(r) + \sum_r gP(r) = \langle f \rangle + \langle g \rangle$$

ale pre súčin funkcií $\langle fg \rangle = \langle f \rangle \langle g \rangle$ to platí len ak funkcie f a g sú nezávislé, inak nie.

2.3 ROZDELENIE HUSTOTY PRAVDEPODOBNOTI

Spojité premenné vyžadujú trochu rozdielne spracovanie ako diskrétné. Pravdepodobnosť, že x leží v špecifikovanom rozsahu je popísaná rozdelením hustoty pravdepodobnosti

Pravdepodobnosť toho, že výsledok leží medzi x_1 a x_2 $= \int_{x_1}^{x_2} P(x)dx$

alebo $P(x) = \lim_{\delta x \rightarrow 0} \frac{\text{pravdepodobnosť (výsledok leží medzi } x \text{ a } x + \delta x)}{\delta x}$

pre očakávané hodnoty $\langle x \rangle = \int_{-\infty}^{\infty} xP(x)dx$

$$\langle f \rangle = \int_{-\infty}^{\infty} f(x)P(x)dx$$

Príklad na rozdelenie pravdepodobnosti

Hádzanie 4 mincí. Vypočítajte pravdepodobnosti jednotlivých možných konfigurácií.

Riešenie

- pravdepodobnosť, že prvá minca po dopade bude mať znak je $1/2$, to isté platí pre ostatné mince, takže pravdepodobnosť, že všetky mince budú mať hore znak je daná súčinom jednotlivých pravdepodobností $P(4) = (1/2)^4 = 1/16$

- Ďalšia možnosť je, že po dopade tri mince budú mať hore znak a jedna minca bude mať hore číslo - pre prvé tri pravdepodobnosť je súčin jednotlivých pravdepodobností $= 3 \times 1/2 = 1/8$ a pre štvrtú je pravdepodobnosť $1/2$. Ak však nezáleží na poradí, ktorá minca bude mať hore číslo, potom máme 4 možnosti (ZZZC, ZZZC, ZCZZ, CZZZ) každá s rovnakou pravdepodobnosťou $1/16$, takže $P(3) = 4 \times 1/16 = 1/4$

- pre dva znaky a dve čísla máme 6 permutácií mincí (ZZCC, ZCZC, ZCCZ, CCZZ, CZCZ, CZZC) každá s pravdep. $1/16$, takže $P(2) = 6/16 = 3/8$

- pre 1 znak a tri čísla je $P(1) = P(3) = 1/4$

- pravdepodobnosť, že nepadne žiadny znak $P(0) = P(4) = 1/16$

pravdepodobnosť, že sa niečo stane musí byť 1,

ak r bude počet hodených znakov ($r=0, 1, 2, 3, 4$), máme pravdepodobnosti $P(r) = (1/16, 1/4, 3/8, 1/4, 1/16)$ toho, že hod štyrmi mincami dá r znakov - toto príklad rozdelenia pravdepodobnosti.

2.3.1 Binomické rozdelenie

Popisuje procesy s daným počtom identických pokusov s dvomi možnými výsledkami. Príkladom binomického rozdelenia je napr. hádzanie mincou, skúška kvality komponentov (prijatý alebo nie), liečenie pacientov (mŕtvy, vyliečený) a pod.

Označme pravdepodobnosť úspechu p , pravdepodobnosť neúspechu potom bude $1-p$, n je počet pokusov, rozdelenie dáva pravdepodobnosť r úspechov a $(n-r)$ neúspechov, mimo týchto n pokusov, každý z ktorých má pravdepodobnosť úspechu p , pravdepodobnosť r úspechov z n pokusov

1) je 2^n možných permutácií úspechu a neúspechu, z ktorých počet s r úspechmi je počet ciest vybrania r z n

$${}_n C_r = \frac{n!}{r!(n-r)!}$$

2) je tu r úspechov s pravdepodobnosťou p a podobne $(n-r)$ neúspechov s pravdepodobnosťou $(1-p)$, kombináciou dostaneme

$$p^r(1-p)^{n-r}$$

Keď spojíme oba faktory, dostaneme **binomické rozdelenie pravdepodobnosti**

$$P(r; p, n) = p^r (1-p)^{n-r} \frac{n!}{r!(n-r)!}$$

pravdepodobnosť závisí nielen na počte úspechov r , ale tiež na vlastnej pravdepodobnosti p a počte pokusov n

Dôležité vlastnosti binomického rozdelenia

Celková pravdepodobnosť že sa niečo stane je

$$\sum_{r=0}^n p^r (1-p)^{n-r} {}_n C_r = [p + (1-p)]^n = 1^n = 1$$

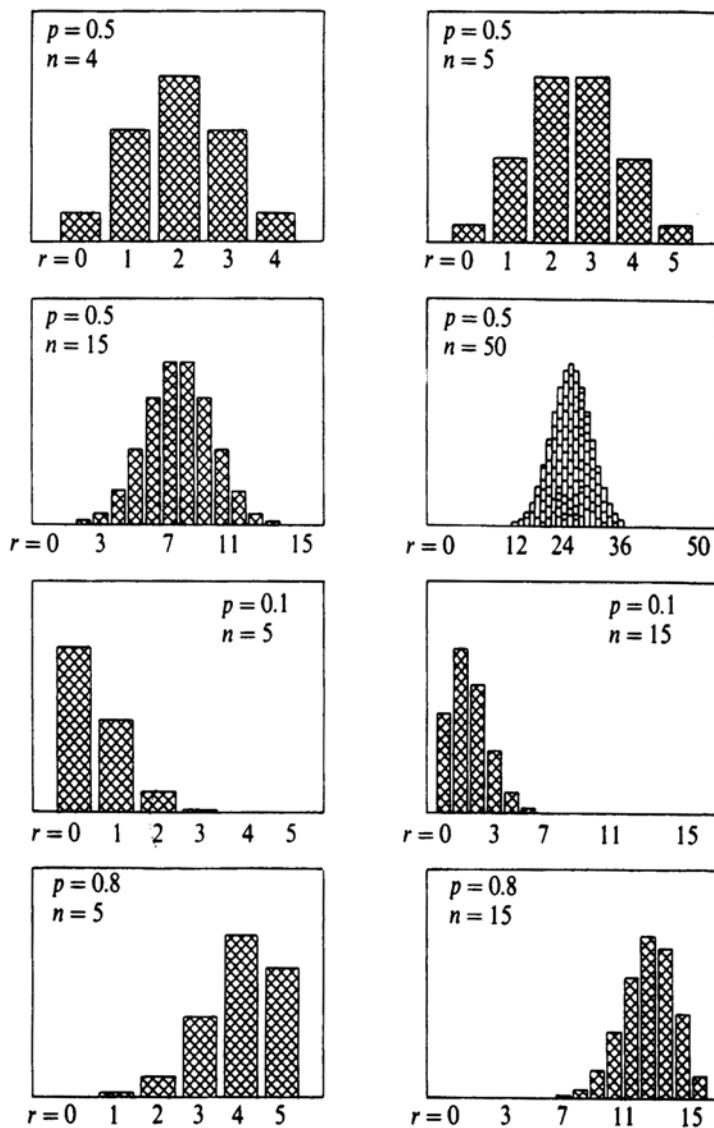
stredný počet úspechov je $\langle r \rangle = np$

rozptyl $V(r) = np(1-p)$

štandardná odchýlka $\sigma = \sqrt{np(1-p)}$

keď p je neznáme môžeme pre rozptyl použiť odhad

$$s^2 = \frac{N}{N-1} n \frac{\bar{r}}{N} \left(1 - \frac{\bar{r}}{N} \right)$$



Obr. 2.2 Binomické rozdelenia s rôznymi hodnotami p a n . Keď $p = 0$, $p = 1$ alebo $p = 0.5$, $\mu = 0$. Vo všeobecnosti rozptyl je menší ako priemer.

Príklad binomického rozdelenia

Hádanie kariet - karty majú 5 rôznych symbolov - t. j. je šanca 20 % uhádnuť správny. Ak hádame 6 kariet, aká je šanca uhádnuť viac než polovicu správne?

Riešenie

Pravdepodobnosť z binomického rozdelenia je $P(4; 0.2, 6) + P(5; 0.2, 6) + P(6; 0.2, 6) = 1.54\% + 0.154\% + 0.0064\% = 1.7\%$.

Pri hádaní šiestich kariet je pravdepodobnosť uhádnuť viac ako polovicu z nich (t.j. najmenej 4 a viac) správne 1.7%.

Binomické rozdelenie je málo používané - častejšie sa stretávame s jeho limitami:

$p \rightarrow 0, n \rightarrow \infty$, ale $np = \text{konšt.}$ binomické \rightarrow Poissonove rozdelenie
 $n \rightarrow \infty, p = \text{konšt.}$, binomické \rightarrow Gaussove rozdelenie

2.3.2 Poissonove rozdelenie

Pri binomickom rozdelení výsledky nastávajú v istom počte pokusov (n). U Poissonovho rozdelenia je to trochu iné - neuvažujeme o počte pokusov, namiesto toho sú tu ostré prípady objavujúce sa v kontinuu (napr. cez búrku bude určitý počet bleskov (ostré prípady), ale nemá zmysel pýtať sa ako často nebol blesk. Geigerov počítač produkuje pri zdroji žiarenia určité praskanie, ale nie určité nepraskanie).

Predpokladajme, že v nejakom intervale priemere nastane λ eventov. Rozdeľme interval do n malých rovnakých častí, takých malých, že šanca mať dva eventy v jednej časti je zanedbateľná. Potom pravdepodobnosť, že daná sekcia obsahuje event je $p = \lambda / n$. Pravdepodobnosť, že bude r eventov v n sekciách intervalu je daná binomickým rozdelením

$$P(r; \lambda / n, n) = p \frac{\lambda^r}{n^r} \left(1 - \frac{\lambda}{n}\right)^{n-r} \frac{n!}{r!(n-r)!}$$

pre $n \rightarrow \infty$ s r konečným - faktoriál dáva

$$\frac{n!}{(n-r)!} = n(n-1)(n-2)\dots(n-r+1) \rightarrow n^r \quad \text{a} \quad \left(1 - \frac{\lambda}{n}\right)^{n-r} \rightarrow \left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda}$$

Poissonove rozdelenie - pravdepodobnosť získania r eventov ak stredný očakávaný počet je λ je

$$P(r; \lambda) = \frac{\lambda^r}{r!} e^{-\lambda}$$

Dôležité vlastnosti Poissonovho rozdelenia -

celková pravdepodobnosť = 1 takže $\sum_{r=0}^{\infty} P(r; \lambda) = 1$

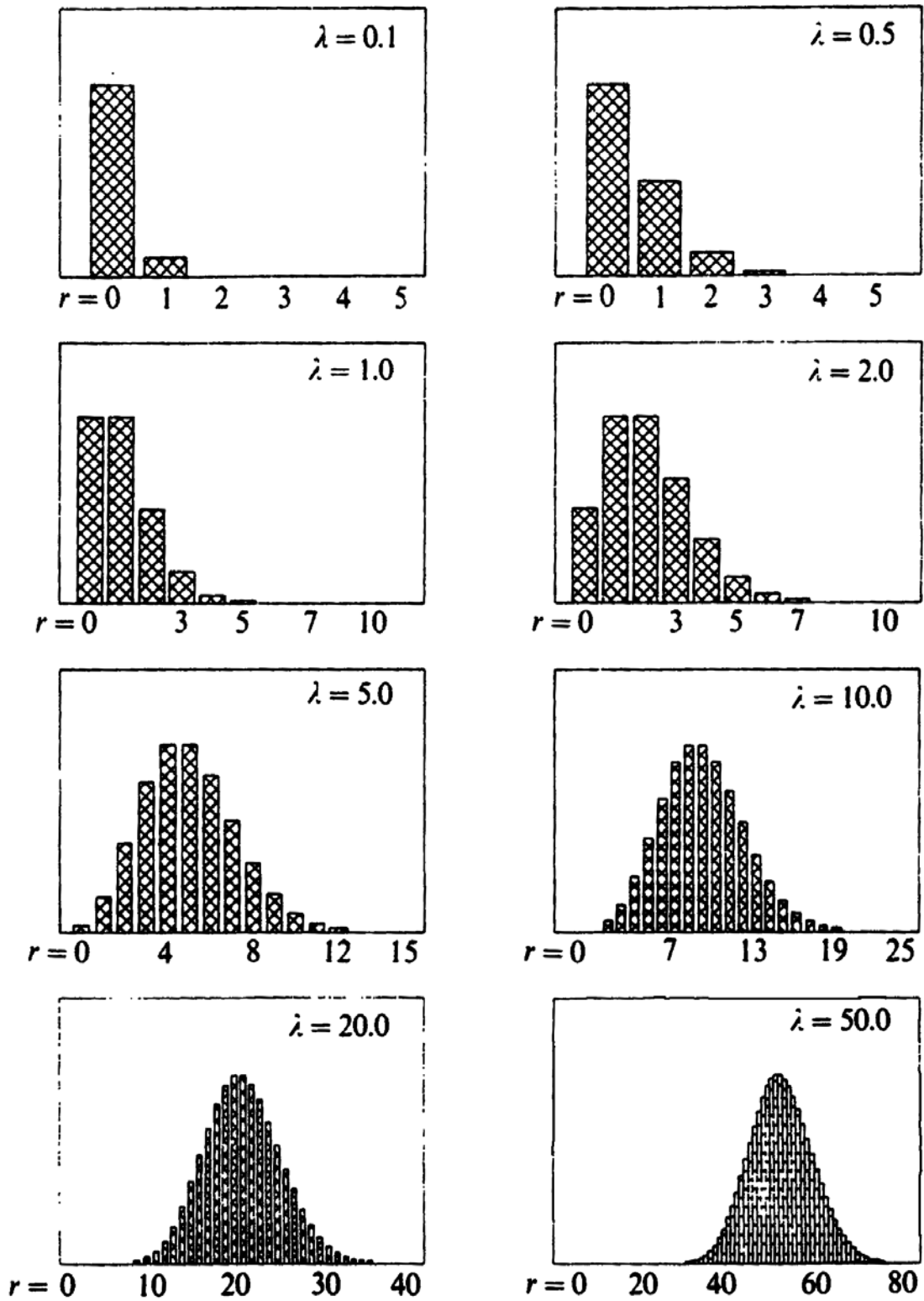
stredná hodnota počtu eventov je $\langle r \rangle = \lambda$

rozptyl $V(r) = \lambda$

štandardná odchýlka $\sigma = \sqrt{\lambda}$

štandardná odchýlka sa rovná druhej odmocnine priemerného počtu eventov.

Z obr 2.3 vidieť, že pre $\lambda < 1.0$ napravdepodobnejší výsledok je 0. Poissonovské rozdelenie je vždy širšie ako binomické s rovnakou strednou hodnotou. Poissonovský rozptyl je rovný priemeru λ , zatiaľ čo binomický rozptyl $np(1-p)$ je vždy menší ako priemer np .



Obr. 2.3 Poissonovské rozdelenia s rôznymi hodnotami λ .

Príklad na Poissonove rozdelenie

Počet neutrín detegovaných v 10 sekundových intervaloch v experimente 23.2.1987 (v čase keď astronómovia prvý raz pozorovali supernovu 1987a)

Počet eventov	: 0	1	2	3	4	5	6	7	8	9
Počet intervalov	: 1042	860	307	78	15	3	0	0	0	1
Predpoveď	: 1064	823	318	82	16	2	0.3	0.03	0.003	0.0003

$$\text{priemer} \quad \bar{r} = \frac{860 + 2 \times 307 + 3 \times 78 + 4 \times 15 + 5 \times 3}{1042 + 860 + 307 + 78 + 15 + 3} = 0.77$$

Poissonova predpoveď dobre súhlasí s dátami (okrem hodnoty pre 9 eventov). To ukazuje, že pozadie spôsobené náhodnými eventami je Poissonovské, a 9 eventov nie je požadovaná fluktuácia, ale prišli zo supernovy

Poissonovské rozdelenie môže dobre aproximovať binomické rozdelenie pokiaľ počet pokusov n je veľký a pravdepodobnosť p je malá

Príklad na porovnanie Poissonovho a binomického rozdelenia

Vypočítajte predpovede pre 100 pokusov s individuálnou pravdepodobnosťou úspechu 2% z binomického a z Poissonovho rozdelenia pre počet úspechov 0-6.

Riešenie

100 pokusov s individuálnou pravdepodobnosťou úspechu 2%, binomické pravdepodobnosti pre počet úspechov sú

r	: 0	1	2	3	4	5	6
P(binomická)	: 13.3%	27.1%	27.3%	18.2%	9.0%	3.5%	1.1%
Poissonove rozdelenie pre priemer 2 dáva pravdepodobnosti							
P(Poisson)	: 13.5%	27.1%	27.1%	18.0%	9.0%	3.6%	1.2%

2.3.3 Dve Poissonove rozdelenia

Predpokladajme eventy dvoch typov - a, b s zodpovedajúcimi priemermi λ_a, λ_b a taktiež poznáme pravdepodobnosť pozorovania r_a, r_b . Celkove r eventov môže byť takých, že všetky sú typu b, alebo sú kombinácia typu a, b atď.

$$P(r) = \sum_{r_a=0}^r P(r_a; \lambda_a) P(r-r_a; \lambda_b) = e^{-\lambda_a} e^{-\lambda_b} \sum \frac{\lambda_a^{r_a} \lambda_b^{r-r_a}}{r_a! (r-r_a)!} =$$
$$= e^{-(\lambda_a + \lambda_b)} \frac{(\lambda_a + \lambda_b)^r}{r!} \sum_{r_a=0}^r \frac{r!}{r_a! (r-r_a)!} \left(\frac{\lambda_a}{\lambda_a + \lambda_b} \right)^{r_a} \left(\frac{\lambda_b}{\lambda_a + \lambda_b} \right)^{r-r_a}$$

Suma je práve binomický rozvoj $\left(\frac{\lambda_a}{\lambda_a + \lambda_b} + \frac{\lambda_b}{\lambda_a + \lambda_b} \right)^r$, čo je rovné 1, takže výsledok je

$$P(r) = e^{-(\lambda_a + \lambda_b)} \frac{(\lambda_a + \lambda_b)^r}{r!}$$

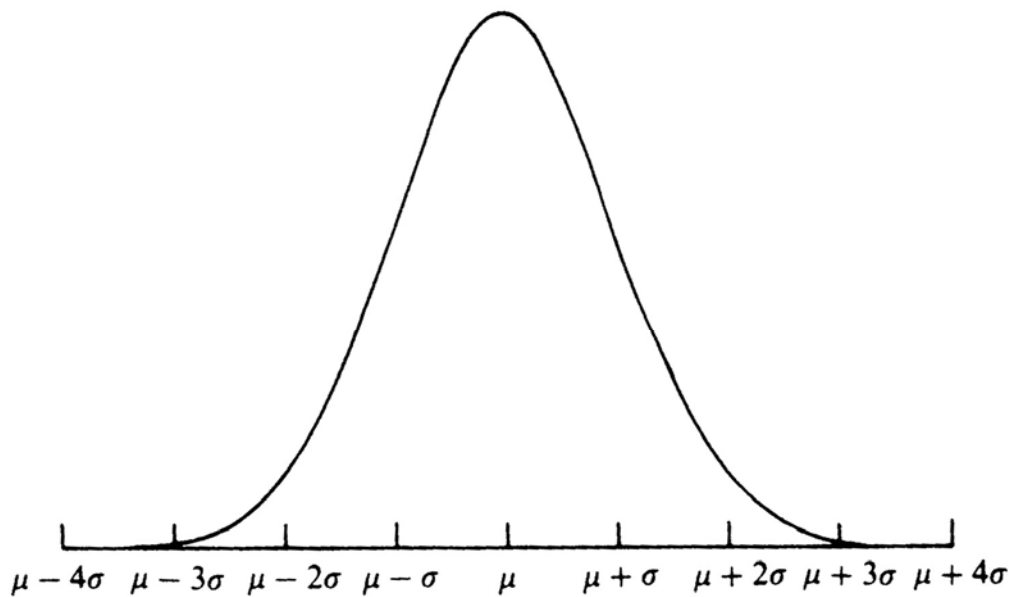
t.j. suma dvoch Poissonovských procesov je ďalší Poissonovský proces.

2.3.4 Gaussovské rozdelenie

Rozdelovacia funkcia hustoty pravdepodobnosti Gaussovho rozdelenia je:

$$P(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Gaussove alebo normálne rozdelenie je symetrické okolo $x = \mu$, šírka je kontrolovaná hodnotou σ (Prvý krát spomínané 1733 de Moivre (Angličan) v práci "Approximatio ad summam terminorum binomii $(a+b)^n$ in seriem expansi"). Pre $x = \mu \pm \sigma$, $P(x)$ padá na 0.61 píkovej hodnoty, t.j. takmer na polovicu. Toto sú inflexné body, kde druhá derivácia je rovná nule.



Obr. 2.4 Gaussovské (normálne) rozdelenie pravdepodobnosti

Ak urobíme substitúciu $z = (x-\mu) / \sigma$ dostaneme $\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ čo sa nazýva jednotkové Gaussove alebo normálne rozdelenie

Dôležité vlastnosti Gaussovho rozdelenia:

je normované na 1

$$\int_{-\infty}^{\infty} P(x; \mu, \sigma) dx = 1$$

μ je stredná hodnota rozdelenia

$$\int_{-\infty}^{\infty} xP(x; \mu, \sigma) dx = \mu$$

je to taktiež modus a medián

Štandardná odchýlka je σ , rozptyl σ^2

$$\int_{-\infty}^{\infty} (x - \mu)^2 P(x; \mu, \sigma) dx = \sigma^2$$

Pre Gaussovské rozdelenie platí - FWHM (šírka v polovici maxima) = 2.35σ ,

Príklad - Výpočet FWHM (v jednotkách σ)

Ukážte, že pre Gaussovské rozdelenie platí - FWHM (šírka v polovici maxima) = 2.35σ .

Riešenie

Keď používame Gaussian je obvykle najjednoduchšie posunúť počiatok súradnicovej sústavy tak, že $\mu = 0$,

$$\frac{f(x)}{2} = \frac{1}{2\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \text{ pre } x=0 \quad \frac{f(0)}{2} = \frac{1}{2\sigma\sqrt{2\pi}}$$

$$\frac{1}{2\sigma\sqrt{2\pi}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

takže v polovičnej výške dostaneme $\frac{1}{2} = e^{-\frac{x^2}{2\sigma^2}}$

$$\ln 1 = \ln 2 - \frac{x^2}{2\sigma^2} \Rightarrow x = \sigma\sqrt{2\ln 2} = 1.1774\sigma$$

Takto sme vypočítali polovicu šírky gausiánu v polovičnej výške, takže plná šírka sa potom rovná $2 \times 1.1774 = 2.3548 \sigma$

Keď používame funkciu gaussovského rozdelenia stretávame sa s rôznymi štandardnými integrálmi. Ich riešenia sa nachádzajú v každej matematickej príručke. Bohužiaľ neurčité integrály gausiánu sa nedajú riešiť analyticky, ale sú numericky spočítané a spracované do tabuliek. V tabulkách sú obvykle uvádzané hodnoty integrovaného gaussovského rozdelenia medzi symetrickými hranicami - $(x-\mu)/\sigma$ a $+(x-\mu)/\sigma$, t.j. pravdepodobnosti, že ak event má Gaussove rozdelenie, bude ležať nejaký počet štandardných odchýliek od strednej hodnoty.

68.27%	oblasti leží vnútri oblasti	1σ
95.45%		2σ
99.73%		3σ

ak chceme vedieť určité % potom

90 %	leží vnútri	1.645σ
95 %		1.960σ
99%		2.576σ
99.9%		3.290σ

Tabelované hodnoty dvojstranného integrálu gausiánu dávajú percentá pravdepodobnosti, že bod leží v danom počte σ od strednej hodnoty (na obr. 2.5 nevyšrafovaná časť)



Obr. 2.5

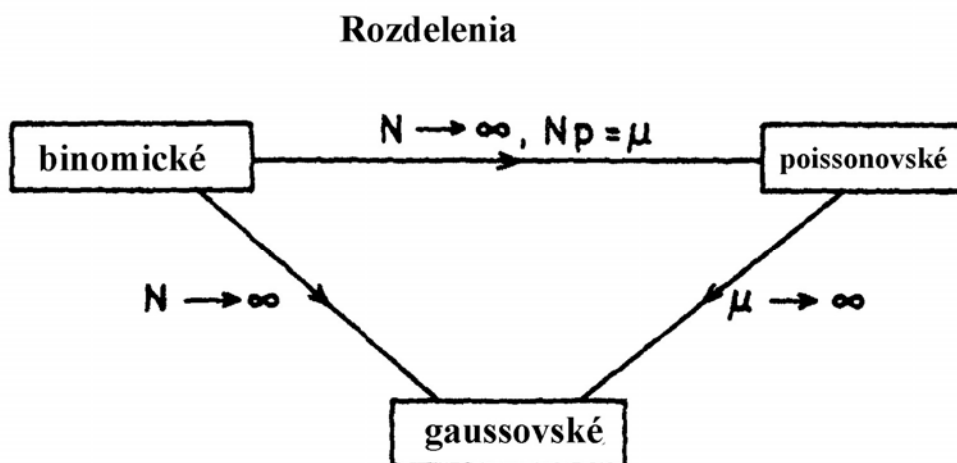


Obr. 2.6

Pokiaľ tabelované hodnoty vyjadrujú jednostranný integrál gaussiánu, udávajú percentá pravdepodobnosti, že bod leží v danom počte štandardných odchýliek σ na jednej strane od strednej hodnoty (na obr.2.6 nevyšrafovaná časť).

2.3.4.1 Gauss ako limita Poissonovského a binomického rozdelenia

Pre veľké hodnoty λ Poissonove rozdelenie prechádza na Gaussove s $\mu = \lambda$, a $\sigma = \sqrt{\lambda}$ (λ by malo byť minimálne 5, istejšie 10. V takom prípade je gaussián veľmi vhodnou aproximáciou poissona).



Obr. 2.7 Vzťahy medzi binomickým, Poissonovým a Gaussovým rozdelením. N a p sú počet pokusov a individuálna pravdepodobnosť úspechu pre binomické rozdelenie. Stredná hodnota pre Poissonove rozdelenie je μ . V prípade Gaussovho rozdelenia odvodeného ako limita Poissonovho rozdelenia pre veľké μ , stredná hodnota μ a rozptyl sú rovnaké.

Príklad: Poissonove rozdelenie aproximované Gaussovým

Porovnajme pravdepodobnosť z Poissonovho rozdelenia pre $\lambda = 5.3$ pre dva alebo menej eventov s pravdepodobnosťou pre rovnaký prípad z Gaussovho rozdelenia.

Riešenie

Ak λ je 5.3, potom pravdepodobnosť dvoch, alebo menej eventov použitím Poissonovho rozdelenia je 10.2 %. Keď aproximujeme histogram Poissonovho rozdelenia hladkou Gaussovou krivkou, zodpovedajúca hodnota pre Gaussián je v polovici medzi možnými diskretnými hodnotami 2 a 3, teda 2.5 eventu. To je $(5.3-2.5)/1.22\sigma$ od priemeru a z tabelovaných hodnôt pre jednostr. Gauss máme pravdepodobnosť 11.1%

Rovnako binomické rozdelenie prechádza na Gauss s $\mu = np$ a $\sigma = \sqrt{np(1-p)}$ p by malo byť ~ 0.5 , väčšie alebo menšie hodnoty p vyžadujú väčšie n , pri väčších početnostiach skoro všetko smeruje ku Gaussu - o tom hovorí **centrálna limitná veta**.

2.3.5 Binormálne rozdelenie

zahŕňa dve premenné x, y , kovariačná matica je $V = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$

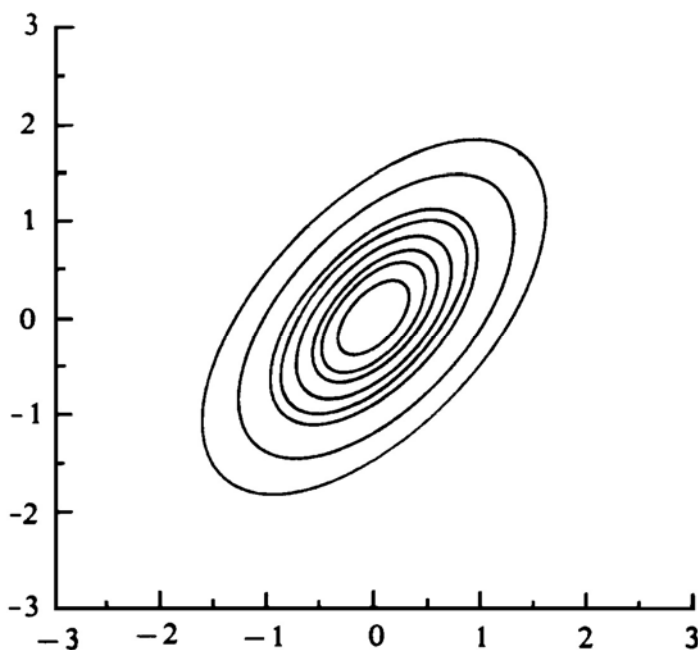
a jej inverzná matica $V^{-1} = \frac{1}{\sigma_x\sigma_y(1-\rho^2)} \begin{pmatrix} \sigma_y^2 & -\rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & \sigma_x^2 \end{pmatrix}$

kde ρ je korelačný koeficient $\rho = \frac{\text{cov}(x, y)}{\sigma_x\sigma_y}$

Výsledný vzťah pre binormálne alebo dvojrozmerné Gaussove rozdelenie je

$$P(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x} \right)^2 + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) \right] \right\}$$

zobrazenie je vrstevnicový graf. Vrstevnice rovnakej pravdepodobnosti sú krivky, pre ktoré exponent v rovnici je konštantný a to je rovnica elipsy. Je vidieť, že elipsa, pre ktorú exponent = - 1/2 má extrémne x a y hodnoty pri $\mu_x \pm \sigma_x$ a $\mu_y \pm \sigma_y$



Obr.2.8 Binormálne rozdelenie

Obr. 2.8 ukazuje línie konštantnej pravdepodobnosti pre dvojrozmerné rozdelenie, kde x a y sú kladne korelované. Elipsy konštantnej pravdepodobnosti sú pri 90, 80, ...

10% pikovej hodnoty. Parametre sú $\sigma_x = \sqrt{\frac{5}{8}}$, $\sigma_y = \sqrt{\frac{5}{8}}$, $\rho = \frac{3}{5}$

Ak spravíme rez rozdelením, uvažujúc rozdelenie y , pre fixnú hodnotu x , potom rovnica dáva Gaussove rozdelenie v y ktorého štandardná odchýlka sa blíži k hodnote $\frac{\sigma_y}{\sqrt{1-\rho^2}}$ a stredná hodnota je $\mu_y + \rho \left(\frac{\sigma_y}{\sigma_x} \right) (x - \mu_x)$

2.4 INÉ ROZDELENIA

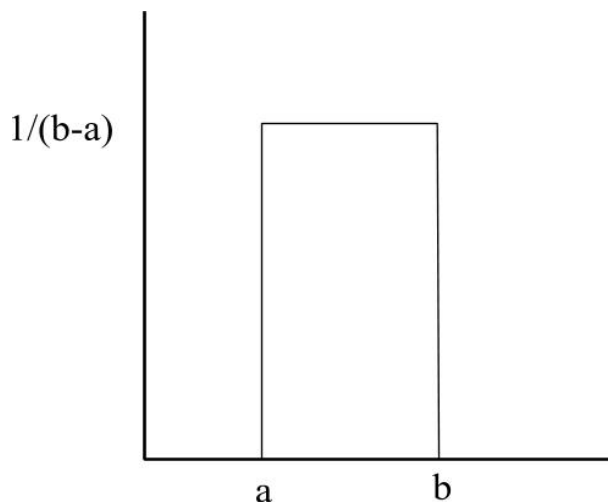
2.4.1 Rovnomerné rozdelenie

$$P(x) = \begin{cases} 1/(b-a) & \text{pre } a \leq x \leq b \\ 0 & \text{pre ostatné} \end{cases}$$

priemer je obvykle $(a+b)/2$, keď spravíme integrál aby sme získali $\langle x^2 \rangle$, potom pre rozptyl dostaneme

$$V(x) = \frac{(b-a)^2}{12}$$

Štandardná odchýlka je odmocnina z rozptylu

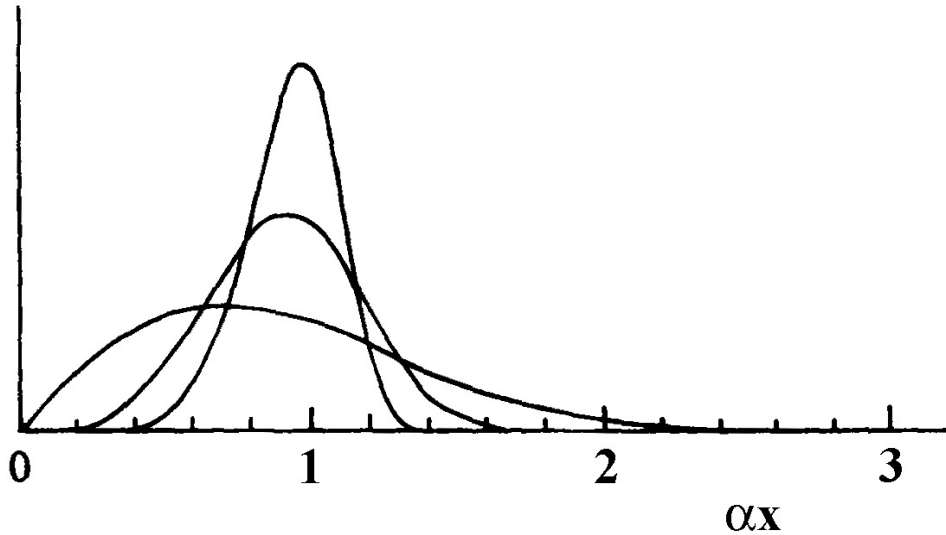


Obr. 2.9 Rovnomerné rozdelenie

2.4.2 Weibullove rozdelenie

$$P(x; \alpha, \beta) = \alpha \beta (\alpha x)^{\beta-1} e^{-(\alpha x)^\beta}$$

α je škálový faktor, β je špicatosť píku, $\beta = 1$ dáva exponenciálnu funkciu



Obr. 2.10 Niekoľko Weibullovych funkcií, píky sú pre $\beta = 2.0, 4.0$ a 7.0

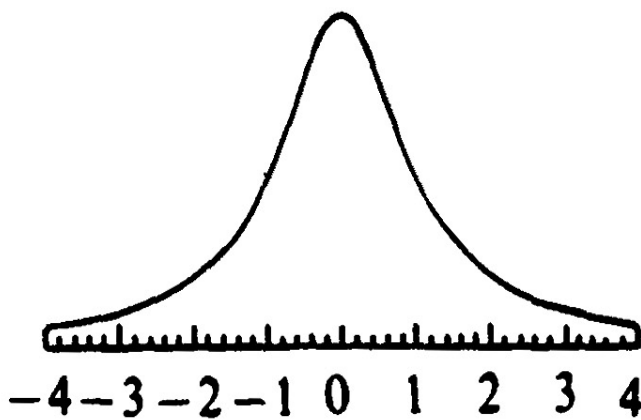
2.4.3 Breit-Wignerove alebo Cauchyho rozdelenie

$$F(m; M, \Gamma) = \frac{1}{2\pi} \frac{\Gamma}{(m - M)^2 + (\Gamma/2)^2}$$

Breit-Wignerova funkcia, často používaná v jadrovej fyzike pre vyjadrenie rozdelenia častíc hmotnosti m v dôsledku rezonancie s hmotnosťou M a šírkou Γ , sa redukuje na Cauchyho funkciu $F(z)$ zmenou počiatku a škály.

$$F(z) = \frac{1}{\pi} \frac{1}{1 + z^2}$$

Cauchyho funkcia nemá rozptyl, lebo integrál $\int z^2 F(z) dz$ diverguje



Obr. 2.11 Cauchyho funkcia

KONTROLNÉ OTÁZKY

18. Čo je to pravdepodobnosť, ako ju môžeme vyjadriť?
19. Ako môžeme vyjadriť pravdepodobnosť, že : a) nastanú súčasne dva nezávislé javy, b) nastane niektorý z dvoch vzájomne sa vylučujúcich javov?
20. O čom hovorí zákon veľkých čísel?
21. Ako vyjadríme očakávanú hodnotu nejakej veličiny?
22. Aká je rozdelovacia funkcia hustoty pravdepodobnosti binomického rozdelenia?
23. Aká je stredná hodnota, rozptyl a štandardná odchýlka pre binomické rozdelenie?
24. Ako môžeme odhadnúť rozptyl pre binomické rozdelenie keď nepoznáme hodnotu p - individuálne pravdepodobnosti úspechu?
25. Aká je rozdelovacia funkcia hustoty pravdepodobnosti Poissonovho rozdelenia?
26. Aká je stredná hodnota, rozptyl a štandardná odchýlka pre Poissonove rozdelenie?
27. Keď sčítame dva Poissonovské procesy aké rozdelenie bude mať súčet?
28. Aká je rozdelovacia funkcia hustoty pravdepodobnosti Gaussovho rozdelenia?
29. Aká je stredná hodnota, rozptyl a štandardná odchýlka pre Gaussove rozdelenie?
30. Aký je vzťah medzi FWHM a σ pre Gaussove rozdelenie?
31. Aká je pravdepodobnosť, že skutočná hodnota veličiny, ktorá má normálne rozdelenie bude ležať v intervale $\pm 1\sigma$ od strednej hodnoty?
32. Kedy prechádza Poissonove rozdelenie na Gaussove?

ÚLOHY

2.1 Počas sprišky meteoritov, meteority padajú s početnosťou 15.7 za hodinu. Aká je pravdepodobnosť pozorovania menej než 5 meteoritov v časovej perióde 30 minút?

2.2 Študent sa pokúša stopnúť auto. Autá prechádzajú v náhodných intervaloch, pri priemernej početnosti 1 za minútu. Pravdepodobnosť, že auto zastaví je 1%. Aká je pravdepodobnosť, že študent bude ešte čakať : a) po prechode 60 aut ? b) po 1 hodine?

2.3 Ukážte, že šikmost' a špicatosť sú pre Gausián nulové.

SÚHRN

Pravdepodobnosť p získania určitých špecifických výsledkov realizáciou jedného z týchto meraní je možné vyjadriť pomerom

$$p = \frac{\text{počet udalostí v ktorých pozorujeme výsledok}}{\text{celkový počet meraní}}$$

vzájomne sa vylučujúce javy - súčet pravdepodobností $p = p_1 + p_2 + \dots + p_N$
nezávislé javy - súčin pravdepodobností $p = p_1 p_2 p_3 \dots p_N$
podmienené pravdepodobnosti - $p = p_A \cdot p_{A|B}$
Očakávaná hodnota r - označuje sa $\langle r \rangle$, alebo niekedy $E(r)$ (expected value).

$$\langle r \rangle = \sum_r r P(r)$$

binomické rozdelenie pravdepodobnosti

$$P(r; p, n) = p^r (1-p)^{n-r} \frac{n!}{r!(n-r)!}$$

pravdepodobnosť závisí nielen na počte úspechov r , ale tiež na vlastnej pravdepodobnosti p a počte pokusov n

stredný počet úspechov je $\langle r \rangle = np$

rozptyl $V(r) = np(1-p)$

štandardná odchýlka $\sigma = \sqrt{np(1-p)}$

keď p je neznáme môžeme pre rozptyl použiť odhad $s^2 = \frac{N}{N-1} n \frac{\bar{r}}{N} \left(1 - \frac{\bar{r}}{N}\right)$

Poissonove rozdelenie -

pravdepodobnosť získania r eventov ak stredný očakávaný počet je λ je

$$P(r; \lambda) = \frac{\lambda^r}{r!} e^{-\lambda}$$

celková pravdepodobnosť = 1 takže $\sum_{r=0}^{\infty} P(r; \lambda) = 1$

stredná hodnota počtu eventov je $\langle r \rangle = \lambda$

rozptyl $V(r) = \lambda$

štandardná odchýlka $\sigma = \sqrt{\lambda}$

suma dvoch Poissonovských procesov je ďalší Poissonovský proces

Rozdelovacia funkcia hustoty pravdepodobnosti Gaussovského rozdelenia je:

$$P(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

je normované na 1 $\int_{-\infty}^{\infty} P(x; \mu, \sigma) dx = 1$

μ je stredná hodnota rozdelenia $\int_{-\infty}^{\infty} xP(x; \mu, \sigma) dx = \mu$

je to taktiež modus a medián

Štandardná odchýlka je σ , rozptyl σ^2 $\int_{-\infty}^{\infty} (x - \mu)^2 P(x; \mu, \sigma) dx = \sigma^2$

Pre Gaussovské rozdelenie platí - FWHM (šírka v polovici maxima) = 2.35σ ,
Rovnomerné rozdelenie

$$P(x) = \begin{cases} \frac{1}{b-a} & \text{pre } a \leq x \leq b \\ 0 & \text{pre ostatné} \end{cases}$$

priemer je obvykle $(a+b)/2$, rozptyl $V(x) = \frac{(b-a)^2}{12}$

3 NEISTOTY (ERRORS)

UČEBNÉ CIELE

V tejto kapitole by sa mal študent oboznámiť s typmi neistôt, naučiť sa pracovať s nimi, oboznámiť sa s centrálnou limitnou vetou, ako sa kombinujú neistoty, čo sú systematické a náhodné neistoty

KLÚČOVÉ SLOVÁ

neistoty - náhodné a systematické, neistoty typu A a typu B, kombinácie neistôt, centrálna limitná veta

3.1 ÚVOD

Pri meraniach sa stretávame s neistotami, ktoré môžu byť systematické alebo náhodné. Neistoty môžu mať mnohé príčiny, tvoria ich viaceré komponenty. Keď hodnoty sú ohodnotené neistotou, táto neistota je gaussovská štandardná odchýlka σ . Ak napr. povieme, že tehla váži 10.8 ± 0.1 kg, myslí sa tým, že to vážime váhou, ktorá dáva hodnotu, ktorá sa v rámci ± 0.1 kg nelíši od skutočnej hodnoty so 68% pravdepodobnosťou, v rámci 0.2 kg s 95% a 0.3 kg s 99.7 % pravdepodobnosťou. Toto nie je len taký náhodný výber. Neistoty v meraniach a

výsledkoch sú vo všeobecnosti dobre popísané Gaussovým rozdelením (preto sa toto rozdelenie nazýva normálnym rozdelením)

3.2 PREČO SÚ NEISTOTY GAUSSOVSKÉ?

Zdroje neistôt pri meraniach môžu byť rôzne. Napríklad ak meriame dĺžku nejakého predmetu pomocou pravítka - výsledok je ovplyvnený najrôznejšími príčinami - uhol pohľadu, kalibrácia pravítka, chvenie rúk a pod. Pri použití iných meradiel, napr. digitálnych - do hry vstupujú iné okolnosti ktoré ovplyvňujú získanú hodnotu. Nedokonalosť merania nemá len jednu príčinu, ale mnohé. V nasledovnom si ukážeme zaujímavé a významné vlastnosti premenných, ktoré sú sumou niekoľkých ďalších. Hovorí o tom centrálna limitná veta (central limit theorem - CLT).

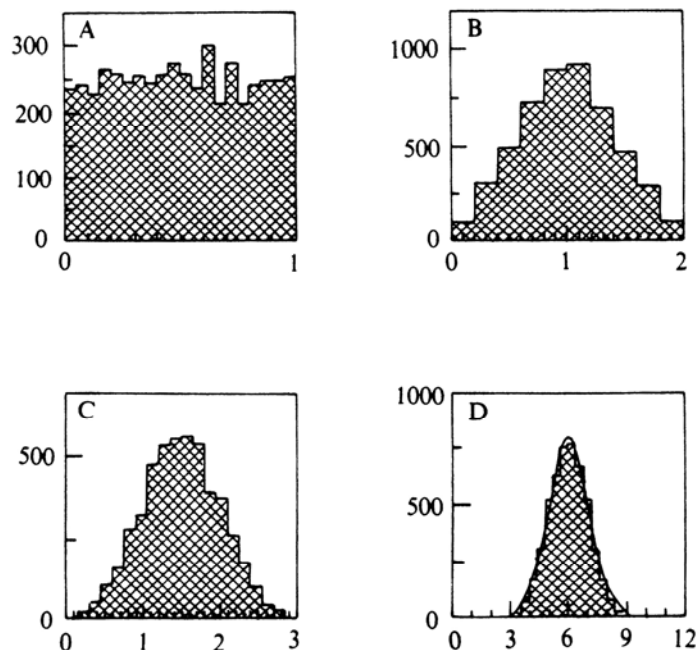
3.2.1 Centrálna limitná veta (central limit theorem - CLT)

Máme sumu X pre N nezávislých premenných x_i , kde $i = 1, 2, \dots, N$, každá má z rozdelenia strednú hodnotu μ_i , rozptyl V_i alebo σ_i^2 , rozdelenie pre X má:

a) očakávanú hodnotu $\langle X \rangle = \sum \mu_i$ b) rozptyl $V(X) = \sum V_i = \sum \sigma_i^2$

c) stáva sa Gaussovským pre $N \rightarrow \infty$.

CLT pracuje lepšie v centre rozdelenia ako ďalej od neho. Rozdelenie môže byť nerozoznateľné od Gaussa v oblasti 1-2 sigma od píku, ale na okrajoch, chvostoch sa môže líšiť.



Obr 3.1. A - histogram 5000 čísel náhodne vybraných z rovnomerného rozdelenia medzi 0 a 1. Má strednú hodnotu 1/2, rozptyl 1/12 a je ploché - určite nie Gauss
 B - histogram iných 5000 čísel, každé je sumou dvoch náhodných čísel (ako pre A) $X = x_1 + x_2$, rozdelenie je trojuholníkové, pík pre 1.0, padá lineárne k 0 a 2.
 C - suma 3 náhodných čísel, pík pri 1.5, tvar krivkový, D - suma 12 náhodných čísel - Gauss - stredná hodnota pre 6.0, rozptyl = $12 \times 1/12 = 1.0$

3.3 PRÁCA S NEISTOTAMI

3.3.1 Opakovanie meraní

Predpokladajme, že veličina je meraná mnohokrát. Môžeme využiť CLT v jednoduchšej podobe za predpokladu všetky stredné hodnoty μ_i majú rovnakú hodnotu μ - aj štandardné odchýlky σ_i majú rovnakú hodnotu σ

$$\langle X \rangle = \sum \mu = N\mu$$

$\langle X \rangle$ je očakávaná stredná hodnota sumy N nezávislých premenných x_i , každá z ktorých má strednú hodnotu rozdelenia μ_i a rozptyl V_i (alebo σ_i^2)

a keď si označíme $\bar{x} = \frac{X}{N}$ potom dostaneme

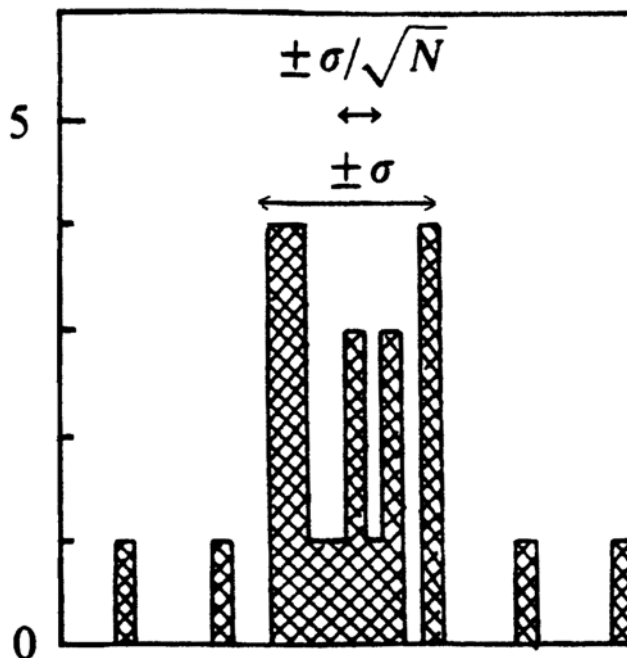
$$\langle \bar{x} \rangle = \mu$$

keď merania sú nezávislé, rozptyl \bar{x} je práve rozptyl X delený N^2

$$V(\bar{x}) = \frac{1}{N^2} \sum V_i = \frac{\sigma^2}{N}$$

(Keď urobíme N meraní x_1, x_2, \dots, x_N a spriemerujeme ich - dostaneme priemer \bar{x} . Tento výsledok je subjektom štatistických fluktuácií, ale v priemere jeho hodnota bude μ , - raz viac, raz menej, ale v priemere μ , t.j. $\langle \bar{x} \rangle = \mu$. Rozdiel medzi aktuálnou hodnotou priemeru \bar{x} a "skutočnou" strednou hodnotou je popísaný rozdelením,

ktoré má rozptyl $V(\bar{x}) = \frac{\sigma^2}{N}$



Obr.3.2 25 meraní - σ štandardná odchýlka priemeru

Štandardná odchýlka priemeru (za spomínaných predpokladov) sa znižuje ako $\frac{1}{\sqrt{N}}$.

Toto je zahrnuté aj v známom štatistickom pravidle: "Priemerovanie je pre teba výhodné".

Ak urobíme N nezávislých meraní niečoho, ich priemer má očakávanú hodnotu ktorá je práve požadovaná veličina, ak jej rozptyl sa znižuje ako $1/N$, rozlíšenie alebo neistota priemeru je $\frac{\sigma}{\sqrt{N}}$, čo je o faktor $\frac{1}{\sqrt{N}}$ menšia neistota ako neistota jednotlivého merania (je dôležité si všimnúť, že toto platí pre každé rozdelenie, nie len pre Gaussovské). $\frac{\sigma}{\sqrt{N}}$ sa nazýva štandardná odchýlka priemeru, popisuje ako

dobre vieme strednú hodnotu rozdelenia, ktorá je často veľmi dôležitou veličinou. Závisí na štandardnej odchýlke rozdelenia a tiež na počte meraní. Obr. 3.2 ukazuje histogram náhodne vzorkovaného Gaussovho rozdelenia - 25 meraní, σ je známa, štandardná odchýlka priemeru týchto 25 meraní je 5x menšia ako σ .

3.3.2 Priemerovanie vážených meraní

Predpokladajme, že máme súbor meraní $\{x_i\}$ nejakej veličiny μ a tieto merania majú rôznu neistotu σ_i . - merania s malou neistotou by mali mať väčšiu váhu ako s veľkou.

Príklad

Vypočítajte priemernú hodnotu ak použijeme 2 voltmetre, jeden má presnosť 0.02V, druhý 0.01V. S prvým voltmetrom nameriame 3.11V, s druhým 3.13V. Akorobiť priemer?

Riešenie

Ak s prvým voltmetrom získame 4 hodnoty, priemer bude mať presnosť $0.02/\sqrt{4} = 0.01V$. 4 merania s horším meracím prístrojom sú ekvivalentné jednému s dvakrát lepšou presnosťou, potom priemerná hodnota nameraného napätia je $V = 1/5 \times 3.11 + 4/5 \times 3.13 = 3.126V$.

Zovšeobecnene, keď sa zaoberáme meraniami rovnakej veličiny z ktorých každá má neistotu σ_i môžeme vyjadriť vážený priemer a jeho rozptyl nasledovne

$$\bar{x} = \frac{\sum x_i / \sigma_i^2}{\sum 1 / \sigma_i^2} \quad V(\bar{x}) = \frac{1}{\sum 1 / \sigma_i^2}$$

Váhovanie výsledkov vyžaduje opatrnosť a uvážlivosť

3.4 KOMBINÁCIE NEISTÔT

3.4.1 Jedna premenná

Predpokladajme, že je f jednoduchá lineárna funkcia $x - f = ax + b$

a, b sú konštanty, x má rozdelenie s rozptylom $V(x)$, x reprezentuje meranie, alebo priebežný výsledok - f - finálny výsledok, alebo tiež ďalší priebežný. Rozptyl výslednej funkcie je:

$$\begin{aligned}
V(f) &= \langle f^2 \rangle + \langle f \rangle^2 = \langle (ax+b)^2 \rangle - \langle ax+b \rangle^2 = \\
&= a^2 \langle x^2 \rangle + 2ab \langle x \rangle + b^2 - a^2 \langle x \rangle^2 - 2ab \langle x \rangle - b^2 = \\
&= a^2 (\langle x^2 \rangle - \langle x \rangle^2) = a^2 V(x)
\end{aligned}$$

alebo inak $\sigma_f = |a| \sigma_x$

b je konštanta jej pridanie k premennej nemá žiadny vplyv na rozmazanie.

zovšeobecnenie - funkciu f(x) pre malé rozdiely môžeme rozvinúť do Taylorovho radu okolo x_0 :

$$f(x) = f(x_0) + (x - x_0) \left(\frac{df}{dx} \right) \Big|_{x=x_0}$$

dosadením do predošlých vzťahov dostaneme

$$V(f) \approx \left(\frac{df}{dx} \right)^2 V(x)$$

$$\sigma_f \approx \left| \frac{df}{dx} \right| \sigma_x$$

aproximácia je platná pre 'malé neistoty' - malé znamená, že prvá derivácia sa nemení viac ako o niekoľko σ .

príklad - napr. - ak rýchlosť náboja je 200 ± 10 m/s, vzdialenosť po 6 sekundách bude 1200 ± 60 m/s

príklad

Aká je neistota $\sin\theta$ ak uhol θ je známy s presnosťou 0.01 radiánu?

Riešenie

Ak uhol θ je známy s presnosťou 0.01 radiánu, potom $\sin\theta$ je známy s presnosťou $0.01|\cos\theta|$

3.4.2 Funkcia dvoch a viacerých premenných

predpokladajme funkciu dvoch premenných, podobne ako v predošlom prípade - lineárnu - v tvare $f(x, y) = ax + by + c$ a, b, c sú konštanty rozvojom, ako v predošlom prípade dostaneme pre rozptyl

$$\begin{aligned}
V(f) &= a^2 (\langle x^2 \rangle - \langle x \rangle^2) + b^2 (\langle y^2 \rangle - \langle y \rangle^2) + 2ab (\langle xy \rangle - \langle x \rangle \langle y \rangle) = \\
&= a^2 V(x) + b^2 V(y) + 2ab \text{cov}(x, y)
\end{aligned}$$

potom pre obecnjšiu $f(x, z)$, opäť pre "malé" neistoty, dostaneme

$$V(f) = \left(\frac{df}{dx} \right)^2 V(x) + \left(\frac{df}{dy} \right)^2 V(y) + 2 \left(\frac{df}{dx} \right) \left(\frac{df}{dy} \right) \text{cov}(x, y)$$

$$\sigma_f^2 = \left(\frac{df}{dx} \right)^2 \sigma_x^2 + \left(\frac{df}{dy} \right)^2 \sigma_y^2 + 2 \left(\frac{df}{dx} \right) \left(\frac{df}{dy} \right) \rho \sigma_x \sigma_y$$

kde diferenciály sú ohodnotené, tak ako predtým, pre skutočné, alebo merané hodnoty (x, z)

3.4.3 Pravidlá kombinácie neistôt

pre x,y nezávislé

$$V(f) = \left(\frac{df}{dx}\right)^2 V(x) + \left(\frac{df}{dy}\right)^2 V(y)$$

$$\sigma_f^2 = \left(\frac{df}{dx}\right)^2 \sigma_x^2 + \left(\frac{df}{dy}\right)^2 \sigma_y^2$$

teda pre funkciu x, y a z (nezávislých) máme

$$\sigma_f^2 = \left(\frac{df}{dx}\right)^2 \sigma_x^2 + \left(\frac{df}{dy}\right)^2 \sigma_y^2 + \left(\frac{df}{dz}\right)^2 \sigma_z^2$$

Príklad

Vypočítajte celkovú neistotu dráhy, ktorú prejde teleso za 6 sekúnd ak sa pohybuje rýchlosťou 200 ± 10 m/s a má zrýchlenie 12 ± 2 m/s².

Riešenie

Ak sa teleso pohybuje rýchlosťou 200 ± 10 m/s a má zrýchlenie 12 ± 2 m/s², potom za 6 sekúnd prejde ($s = vt + 1/2 at^2$) vzdialenosť 1416 m, celková neistota - príspevok rýchlosti je ± 60 m, príspevok zrýchlenia ± 36 m, potom výsledok je ± 70 m

3.4.4 Relatívne neistoty

ak máme $f = x y$ podľa predošlého pre rozptyl dostaneme

$$V(f) = y^2 V(x) + x^2 V(y), \quad \text{čo môžeme rozpísať ako} \quad \left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$$

ak máme $f = \frac{x}{y}$, tak tiež pre podiel dostaneme po úpravách rovnaký vzťah ako

$$\text{pre súčin} \quad \left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2 \quad (*)$$

Príklad

Vypočítajte relatívnu neistotu x.y a x/y ak vieme, že relatívna neistota x je 3% a y 4%

Riešenie

Zo vzťahu (*) pre relatívnu neistotu xy a x/y dostaneme 5%

Percentuálna neistota je výhodná tiež pre prevrátné hodnoty x a je rovnaká ako pre samotné x

$$\frac{\sigma_{1/x}}{1/x} = \frac{\sigma_x}{x}$$

pre logaritmus je relatívna neistota x ako relatívna neistota x $\sigma_{\ln x} = \frac{\sigma_x}{x}$

Príklad

Vyjadrite relatívnu neistotu pre prúd z Ohmovho zákona

Riešenie

Ohmov zákon - prúd vytvorený napätím $V \pm \sigma_V$ na odpore $R \pm \sigma_R$ je $I = \frac{V}{R}$ s relatívnou

neistotou danou ako $\left(\frac{\sigma_I}{I}\right)^2 = \left(\frac{\sigma_V}{V}\right)^2 + \left(\frac{\sigma_R}{R}\right)^2$

3.4.5 Niekoľko funkcií niekoľkých premenných

Predpokladajme m rôznych funkcií $f_1, f_2, f_3, \dots, f_m$ s n rôznymi premennými $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, f_k bude korelovaná s každou inou (aj keď $x_{(i)}$ nie sú) pretože rôzne f_k zdieľajú rovnaké $x_{(i)}$, rozptyly na f_k sú dané ako $V(f_i) = \langle f_i^2 \rangle - \langle f_i \rangle^2$

kovariancia medzi dvomi premennými je definovaná ako $\text{cov}(x_{(i)}, x_{(j)}) = \overline{x_{(i)}x_{(j)}} - \overline{x_{(i)}} \overline{x_{(j)}}$
Kovariancia medzi dvoma premennými v distribučnej funkcii rozdelenia pravdepodobnosti $P(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ je presne ekvivalentná

$\text{cov}(x_{(i)}, x_{(j)}) = \langle (x_{(i)} - \mu_i)(x_{(j)} - \mu_j) \rangle = \langle x_{(i)}x_{(j)} \rangle - \mu_i\mu_j$, kde $\mu_i = \langle x_{(i)} \rangle$. Toto sú elementy kovariančnej matice \mathbf{V} , niekedy tiež nazývanej chybová matica (error matrix), kde jej elementy sú $V_{ij} = \text{cov}(x_{(i)}, x_{(j)})$. Diagonálne elementy tejto matice sú práve rozptyly $V_{ii} = \text{cov}(x_{(i)}, x_{(i)}) = V(x_{(i)}) = \sigma_i^2$. Korelačná matica je jej bezrozmerný ekvivalent

$\rho_{ij} = \frac{\text{cov}(x_{(i)}, x_{(j)})}{\sigma_i\sigma_j}$. Jej elementy musia ležať medzi -1 a 1 a udávajú mieru

korelovanosti dvoch premenných

rozvojom funkcie f_i do Taylorovho radu dostaneme

$$f_i \approx f_i(\mu_1, \mu_2, \dots) + \left(\frac{\partial f_i}{\partial x_1}\right)(x_1 - \mu_1) + \left(\frac{\partial f_i}{\partial x_2}\right)(x_2 - \mu_2) + \dots$$

keď toto použijeme vo vzťahu pre rozptyl potom

$$\begin{aligned} V(f_i) &= \left(\frac{\partial f_i}{\partial x_1}\right)^2 \langle (x_1 - \mu_1)^2 \rangle + \dots + 2\left(\frac{\partial f_i}{\partial x_1}\right)\left(\frac{\partial f_i}{\partial x_2}\right) \langle (x_1 - \mu_1)(x_2 - \mu_2) \rangle + \dots = \\ &= \sum_j \left(\frac{\partial f_i}{\partial x_j}\right)^2 V(x_j) + \sum_j \sum_{k \neq j} \left(\frac{\partial f_i}{\partial x_j}\right)\left(\frac{\partial f_i}{\partial x_k}\right) \text{cov}(x_j, x_k) \end{aligned}$$

čo je trochu zovšeobecnený vzťah pre kombináciu neistôt. Kovarianciu môžeme

vyjadriť ako $\text{cov}(f_k, f_l) = \sum_i \sum_j \left(\frac{\partial f_k}{\partial x_i}\right)\left(\frac{\partial f_l}{\partial x_j}\right) \text{cov}(x_i, x_j)$ pre $\text{cov}(f_i, f_i) = V(f_i)$

ak označíme $G_{ki} = \left(\frac{\partial f_k}{\partial x_i}\right)$ a \mathbf{V}_x a \mathbf{V}_f sú chybové matice pre x a f - môžeme písať

$$\mathbf{V}_f = \mathbf{G} \mathbf{V}_x \overline{\mathbf{G}}$$

\mathbf{V}_x a \mathbf{V}_f sú symetrické štvorcové matice rozmerov $n \times n$ a $m \times m$, \mathbf{G} je obdĺžniková matica $m \times n$

napr. 3 premenné x, y, z , ktoré sú nekorelované a merané s neistotami $\sigma_x, \sigma_y, \sigma_z$

$$\mathbf{G} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

$$\mathbf{V}_f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} = \left(\frac{\partial f}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y} \right)^2 \sigma_y^2 + \left(\frac{\partial f}{\partial z} \right)^2 \sigma_z^2$$

3.4.6 Systematické neistoty

Ak máme merací prístroj, ktorý má neistotu 1% - raz nameriame viac, raz menej - opakovanie meraní a priemerovanie zvýši presnosť. Ale ak merací prístroj systematicky meria napr. o 1% viac - opakovanie meraní nepomôže, chyby nie sú nezávislé, pôsobia spolu, štandardnými štatistickými metódami ich nie je možné eliminovať.

Rôzne nezávislé systematické neistoty, v dôsledku ich nezávislosti, sa skladajú aditívne v kvadráte. Väčšia sa stáva dominantná. Pretože náhodné a systematické neistoty sú vzájomne nezávislé - tiež sa skladajú v kvadráte.

Často sa uvádzajú osobitne napr. $A = -10 \pm 1.2 \pm 2.3$ - prvá náhodná, druhá systematická

3.4.6.1 Používanie systematickej neistoty

Predpokladajme dve merania x_1 a x_2 , ktoré majú systematickú neistotu S a náhodné σ_1 a σ_2 . Môžeme uvažovať x_1 zložené z dvoch častí - náhodnej x_1^R s náhodnou neistotou σ_1 a systematickej x_1^S so systematickou S , to isté je aj pre x_2 - x_1^R a x_2^R sú vzájomne nezávislé a sú nezávislé aj s x_1^S a x_2^S , zatiaľ čo x_1^S a x_2^S sú absolútne korelované.

Rozptyl x_1 potom je $V(x_1) = \langle x_1^2 \rangle - \langle x_1 \rangle^2 = \langle (x_1^R + x_1^S)^2 \rangle - \langle x_1^R + x_1^S \rangle^2 = \sigma_1^2 + S^2$

Podobne $V(x_2) = \sigma_2^2 + S^2$ a kovariancia

$$\text{cov}(x_1, x_2) = \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle = \langle (x_1^R + x_1^S)(x_2^R + x_2^S) \rangle - \langle x_1^R + x_1^S \rangle \langle x_2^R + x_2^S \rangle$$

všetko vypadne, zostanú len absolútne korelované x_1^S a x_2^S a $\text{cov}(x_1, x_2) = \text{cov}(x_1^S, x_2^S) = S^2$ takže rozptylová matica pre x_1 a x_2 má náhodné a systematické neistoty skladané v kvadráte na diagonále

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 + S^2 & S^2 \\ S^2 & \sigma_2^2 + S^2 \end{pmatrix}$$

uvažovali sme konštantnú systematickú neistotu - ak nie je konštantná a je umerná meraniu ($S = \varepsilon x$), potom

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 + \varepsilon^2 x_1^2 & \varepsilon^2 x_1 x_2 \\ \varepsilon^2 x_1 x_2 & \sigma_2^2 + \varepsilon^2 x_2^2 \end{pmatrix}$$

iný príklad - ak máme tri merania x_1 a x_2 , x_3 ktoré majú systematickú neistotu S , náhodné neistoty σ_1 , σ_2 , σ_3 a inú nezávislú systematickú neistotu T , ktorú zdieľajú len x_1 a x_2 ale nie x_3 . Kovariančnú maticu potom môžeme napísať

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 + S^2 + T^2 & S^2 + T^2 & S^2 \\ S^2 + T^2 & \sigma_2^2 + S^2 + T^2 & S^2 \\ S^2 & S^2 & \sigma_3^2 + S^2 + T^2 \end{pmatrix}$$

napr. Ohmov zákon - prúd I je určený meraním napätia V , použitím voltmetra s rozlíšením σ_V , na štandardnom rezistore $R \pm \sigma_R$. Neistota I je daná

$\frac{\sigma_I^2}{I^2} = \frac{\sigma_V^2}{V^2} + \frac{\sigma_R^2}{R^2}$, čo je bez problémov. Ak prístrojom meriame dva prúdy I_1 a I_2 , napätia

V_1 a V_2 majú nezávislé neistoty σ_V , ale odpor je rovnaký pre obe napätia, neistota na ňom je tak systematická - premenujeme ju na S_R

Pre neistoty oboch prúdov sa nič nezmení podľa uvedeného vzťahu, je tu ale nenulová kovariancia medzi obidvomi veličinami $\text{cov}(I_1, I_2) = \frac{\partial I_1}{\partial R} \frac{\partial I_2}{\partial R} S_R^2 = \frac{I_1 I_2}{R^2} S_R^2$

napr. pre rozptyl $I_1 - I_2$ dostaneme

$$\frac{2\sigma_V^2 + (I_1 - I_2)^2 \sigma_R^2}{R^2}$$

KONTROLNÉ OTÁZKY

33. O čom hovorí centrálna limitná veta?
34. Aký je rozptyl priemeru hodnôt s rovnakou štandardnou odchýlkou?
35. Ako môžeme vyjadriť vážený priemer a jeho rozptyl?
36. Aké sú pravidlá kombinácie neistôt nezávislých veličín?
37. Aké sú pravidlá kombinácie neistôt závislých (korelovaných) veličín?
38. Ako sa prejavujú systematické neistoty?

ÚLOHY

3.1 Čo je výhodnejšie - súbor desiatich meraní s rozlíšením 1mm, alebo jedno meranie s rozlíšením 0.2 mm?

3.2 Nájdite najlepší kombinovaný výsledok pre 3 nezávislé merania rýchlosti svetla c :

$$\begin{aligned} &299\,798\,000 \pm 5000 \text{ m/s} \\ &299\,789\,000 \pm 4000 \text{ m/s} \\ &299\,797\,000 \pm 8000 \text{ m/s} \end{aligned}$$

3.3 Ak napätie je určené meraním prúdu 1120 ± 10 mA cez odpor $1400 \pm 30 \Omega$. Aká je jeho hodnota a neistota?

3.4 Ak je určovaný prúd meraním napätia 45 ± 1 V na odpore $900 \pm 10 \Omega$. Aká je jeho hodnota a neistota?

SÚHRN

Centrálna limitná veta

Máme sumu X pre N nezávislých premenných x_i , kde $i = 1, 2, \dots, N$, každá má z rozdelenia strednú hodnotu μ_i , rozptyl V_i alebo σ_i^2 , rozdelenie pre X má:

a) očakávanú hodnotu $\langle X \rangle = \sum \mu_i$

b) rozptyl $V(X) = \sum V_i = \sum \sigma_i^2$

c) stáva sa Gaussovským pre $N \rightarrow \infty$.

Opakovanie meraní

ak merania sú nezávislé, rozptyl \bar{x} je práve rozptyl X delený N^2

$$V(\bar{x}) = \frac{1}{N^2} \sum V_i = \frac{\sigma^2}{N}$$

Pravidlá kombinácie neistôt

funkcia dvoch premenných v tvare - $f(x, y) = ax + by + c$ a, b, c sú konštanty

$$V(f) = \left(\frac{df}{dx}\right)^2 V(x) + \left(\frac{df}{dy}\right)^2 V(y) + 2\left(\frac{df}{dx}\right)\left(\frac{df}{dy}\right) \text{cov}(x, y)$$

$$\sigma_f^2 = \left(\frac{df}{dx}\right)^2 \sigma_x^2 + \left(\frac{df}{dy}\right)^2 \sigma_y^2 + 2\left(\frac{df}{dx}\right)\left(\frac{df}{dy}\right) \rho \sigma_x \sigma_y$$

pre funkciu x, y a z (nezávislých) máme

$$\sigma_f^2 = \left(\frac{df}{dx}\right)^2 \sigma_x^2 + \left(\frac{df}{dy}\right)^2 \sigma_y^2 + \left(\frac{df}{dz}\right)^2 \sigma_z^2$$

Relatívne neistoty

ak máme $f = x y$ alebo $f = x / y$ dostaneme

$$\left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$$

Niekoľko funkcií niekoľkých premenných

Predpokladajme m rôznych funkcií $f_1, f_2, f_3, \dots, f_m$ s n rôznymi premennými $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, f_k bude korelovaná s každou inou (aj keď $x_{(i)}$ nie sú) pretože rôzne f_k zdieľajú rovnaké $x_{(i)}$,

$$V(f_i) = \sum_j \left(\frac{\partial f_i}{\partial x_j}\right)^2 V(x_j) + \sum_j \sum_{k \neq j} \left(\frac{\partial f_i}{\partial x_j}\right) \left(\frac{\partial f_i}{\partial x_k}\right) \text{cov}(x_j, x_k)$$

ak označíme $G_{ki} = \left(\frac{\partial f_k}{\partial x_i}\right)$ a \mathbf{V}_x a \mathbf{V}_f sú chybové matice pre x a f - môžeme písať

$$\mathbf{V}_f = \mathbf{G} \mathbf{V}_x \bar{\mathbf{G}}$$

\mathbf{V}_x a \mathbf{V}_f sú symetrické štvorcové matice rozmerov $n \times n$ a $m \times m$, \mathbf{G} je obdĺžniková matica $m \times n$

4 ODHADY (ESTIMATION)

UČEBNÉ CIELE

V tejto kapitole by sa mal študent oboznámiť s veľmi užitočnými a často používanými pojmami a metódami odhadov a naučiť sa ich používať. Mal by sa bližšie zoznámiť s tým, čo sú to estimátory, ich vlastnosťami, s funkciou vierohodnosti, s metódou maximálnej vierohodnosti, s metódou najmenších štvorcov, ktorej pre jej význam je venovaná samostatná kapitola, s metódou momentov, s tým, čo je to stratifikované vzorkovanie.

KLÚČOVÉ SLOVÁ

estimátor, vychýlenosť, konzistentnosť, funkcia vierohodnosti, maximálna vierohodnosť, metóda momentov, metóda najmenších štvorcov, stratifikované vzorkovanie.

ÚVOD

S odhadmi sa v bežnom živote stretávame veľmi často tam, kde z rôznych príčin nedokážeme získať presné výsledky. Odhady poskytujú nepresne výsledky, ale umožňujú ohodnotiť aj to čo nemôžeme zmerať

V štatistike sú odhady technický pojem, je to procedúra vedúca k výsledkom, ktoré môžu byť nepresné, ale rozsah ich nepresnosti je známy (nie je to aproximácia). Jedna z veľmi často používaných metód je metóda najmenších štvorcov. Pre jej význam a rozsah používania je jej venovaná samostatná kapitola

4.1 ESTIMÁTORY

Estimátor je procedúra ktorú keď aplikujeme na vzorku dát dáva číselnú hodnotu pre charakteristiku zdrojového rozdelenia - inak - charakteristiku alebo parameter zdrojovej rozdelovacej funkcie (napr. μ a σ pre Gauss)

Dobry estimátor musí byť : - **konzistentný** ,
- **nevychýlený**
- **účinný**

predpokladajme, že veličina ktorú meriame je a , jej estimátor označujeme striedkou nad príslušnú veličinu \hat{a} .

Keď \hat{a} aplikujeme na N meraní, toto vytvorí odhad veličiny a , odhad sa líši od skutočnej hodnoty v dôsledku štatistických fluktuácií. pre veľké N - fluktuácie klesajú - nezávislé merania bez systematických chýb. Pre dobrý estimátor rozdiely medzi odhadom a skutočnou hodnotou pre veľké vzorky zanikajú - takýto estimátor je konzistentný estimátor

Estimátor je konzistentný ak smeruje k skutočnej hodnote ak dáta idú do nekonečna

$$\lim_{N \rightarrow \infty} \hat{a} = a$$

pre nekonečné N dostaneme skutočnú hodnotu a , pre menšie N pozorujeme odchýlky od skutočnej hodnoty. V prípade ak nadhodnotenie je rovnako pravdepodobné ako podhodnotenie takýto estimátor voláme nevychýlený (**unbiased**) estimátor

Estimátor je nevychýlený ak jeho očakávaná hodnota je rovná skutočnej hodnote

$$\langle \hat{a} \rangle = a$$

Ak odhad je dobrým meraním skutočnej hodnoty - rozmazanie možných hodnôt minimálne. Takýto estimátor je účinný estimátor

Estimátor je účinný, ak jeho rozptyl je malý

Výber estimátora pre použitie v určitej aplikácii vyžaduje posúdenie - neexistuje ideálny "najlepší" estimátor môžu byť rôzne pre rôzne problémy

4.1.1 Funkcia vierohodnosti

máme určitý súbor dát $\{x_1, x_2, x_3, \dots, x_N\}$ na neho použijeme estimátor \hat{a} pre veličinu a , získame číselnú hodnotu odhadu, ktorá je blízko reálnej hodnoty a závisí na súbore dát.

Keď uvažujeme vlastnosti estimátorov - ideme z druhej strany - miesto hľadania v určitej vzorke dát - uvažujeme známe rozdelenie. Hodnoty dát x_i sú z nejakej funkcie hustoty pravdepodobnosti $P(x_i; a)$, ktorá závisí na a . Tvar funkcie P je daný - a určené (predpokladáme, že a je parameter) \hat{a} je funkcia x_i - má očakávanú hodnotu - môžeme teda povedať - Ak zoberieme vzorku N hodnôt z tohto rozdelenia - z nich vytvoríme (x_1, x_2, \dots, x_N) potom v priemere (na mnohých takých vzorkách) hodnota \hat{a} bude $\langle \hat{a} \rangle$ pravdepodobnosť určitého súboru dát $\{x_1, x_2, x_3, \dots, x_N\}$ je súčin individuálnych pravdepodobností - tento súčin sa volá vierohodnosť $L(x_1, x_2, x_3, \dots, x_N; a)$ je kombinovanou pravdepodobnosťou alebo hustotou pravdepodobnosti ktorá tento určitý súbor x_i bude vytvárať z tejto hodnoty a

$$L(x_1, x_2, x_3, \dots, x_N; a) = P(x_1; a)P(x_2; a) \dots P(x_N; a) = \prod P(x_i; a)$$

Očakávaná hodnota pre každú funkciu z $\{x_i\}$ sa nájde integrovaním cez všetky možné hodnoty všetkých x_i , váhovaním celkovou pravdepodobnosťou, t.j vierohodnosťou

$$\langle f(x_1, x_2, \dots, x_N) \rangle = \int \dots \int f(x_1, \dots, x_N) L(x_1, x_2, \dots, x_N; a) dx_1 \dots dx_N$$

v obvyklom označovaní $\langle f(x_1, \dots, x_N) \rangle = \int f L dX$

platí $\langle \hat{a} \rangle = \int \hat{a} L dX$ $\langle \hat{a}^2 \rangle = \int \hat{a}^2 L dX$

požiadavka konzistencie je $\lim_{N \rightarrow \infty} \langle \hat{a} - a \rangle = 0$

hranica minimálneho rozptylu (MVB- minimum variance bound)

pre nevychýlený estimátor je $V(\hat{a}) \geq \frac{1}{\left\langle \left(\frac{d \ln L}{da} \right)^2 \right\rangle}$ alebo inak $V(\hat{a}) \geq \frac{-1}{\left\langle \left(\frac{d^2 \ln L}{da^2} \right) \right\rangle}$

ak pre nejaký estimátor \hat{a} - $V(\hat{a})$ je rovné MVB potom je estimátor \hat{a} účinný, ak nie - jeho účinnosť je $MVB/V(\hat{a})$

pre Gaussovské rozdelenie vzorkový priemer je účinný estimátor μ

pravdepodobnosť určitého x_i je $p(x_i; \mu) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_i - \mu)^2 / 2\sigma^2}$

dáva celkový logaritmus vierohodnosti $\ln L = -\sum_i \frac{(x_i - \mu)}{2\sigma^2} - N \ln(\sigma\sqrt{2\pi})$

dvojnásobným derivovaním podľa μ dostaneme $\frac{d^2 \ln L}{d\mu^2} = -\frac{N}{\sigma^2}$

vzťah neobsahuje x takže je rovný priamo očakávanej hodnote. Po dosadení dostaneme $MVB = \frac{\sigma^2}{N}$, čo je tiež, podľa centrálnej limitnej vety, rozptyl estimátora

napr. rovnomerné rozdelenie - medzi dvomi hranicami (neznámymi). Vzorkový priemer poskytuje konzistentný a nevychýlený estimátor, ale nie je účinný predpokladajme napr. štyri merané hodnoty 1.1, 1.3, 1.7 a 1.6. Piata hodnota 1.4 bude meniť priemer - a predsa nedá novú užitočnú informáciu. Dve extrémne merania majú len jeden reálny význam - získame stredný rozsah $\hat{\mu} = \frac{1}{2}[\min(x_i) + \max(x_i)]$

estimátor, ktorý pre rovnomerné rozdelenie s šírkou W má rozptyl $W^2/[2(N+1)(N+2)]$. Rozptyl priemeru však je σ^2/N po dosadení zo vzťahu pre rovnomerné rozdelenie dostaneme $W^2/12N$. Teda tu je rozptyl stredného rozsahu menší ako rozptyl priemeru (ak $N > 2$). Tu pre veľké N rozptyl padá ako $1/N^2$ a štandardná odchýlka ako $1/N$ a nie ako obvykle $1/\sqrt{N}$ u náhodných neistôt.

4.1.2 Niektoré základné estimátory

4.1.2.1 Odhad priemeru

priemer reprezentuje najlepší odhad skutočnej strednej hodnoty

$$\hat{\mu} = \bar{x}$$

ako zaručuje centrálna limitná veta - tento odhad je konzistentný a nevychýlený a

jeho rozptyl je $V(\hat{\mu}) = \frac{\sigma^2}{N}$,

N je počet dát, σ štandardná odchýlka zdrojového rozdelenia

4.1.2.2 Odhad rozptylu

uvažujme najprv prípad, keď μ je známe (nie veľmi obvyklý prípad)

obvyklý estimátor je $V(\hat{x}) = \frac{1}{N} \sum (x_i - \mu)^2$

platí $\langle V(\hat{x}) \rangle = \frac{N \langle (x - \mu)^2 \rangle}{N} = \langle (x - \mu)^2 \rangle = V(x)$

keď μ nie je známe - μ nahradíme odhadom $\hat{\mu} = \bar{x}$

$$\widehat{V(x)} = \frac{1}{N} \sum (x_i - \mu)^2 = \frac{1}{N} \sum (x_i^2 - \bar{x}^2)$$

ale toto je vychýlené, Keď vezneme očakávanú hodnotu, dostaneme

$$\langle \widehat{V(x)} \rangle = \frac{N \langle (x^2 - \bar{x}^2) \rangle}{N} = \langle x^2 \rangle - \langle \bar{x}^2 \rangle$$

$\langle x \rangle = \langle \bar{x} \rangle$,potom z centrálnej limitnej vety dostaneme

$$\langle \widehat{V(x)} \rangle = \langle x^2 \rangle - \langle x \rangle^2 - (\langle \bar{x}^2 \rangle - \langle \bar{x} \rangle^2) = V(x) - V(\bar{x})$$

a tiež z CLT platí $V(\bar{x}) = V(x)/N$, takže

$$\langle \widehat{V(x)} \rangle = \left(1 - \frac{1}{N}\right) V(x) = \frac{N-1}{N} V(x) \neq V(x)$$

Dostávame odhad rozptylu zdrojového rozdelenia, ktorý je vychýlený (ale vychýlenie padá ako 1/N a môže byť zanedbané ak N je dostatočne veľké).

Ak sme našli vychýlenosť je ľahko ohodnotiť korekciu na ňu. Ak násobíme N/N-1 (čo je známe ako Besselova korekcia), presne kompenzujeme chýbajúcu vychýlenosť. Potom

$$\widehat{V(x)} = s^2 = \frac{1}{N-1} \sum (x_i - \bar{x})^2$$

je konzistentné a bez vychýlenosti. s^2 je **nevychýlený** odhad rozptylu. Estimátor nepracuje pre N=1

Aký je rozptyl tohoto estimátora - po úpravách a pre veľké N dostaneme

$$V(\widehat{V(x)}) = \left[\langle (x - \mu)^4 \rangle - \langle (x - \mu)^2 \rangle^2 \right]$$

potom pre Gaussovské rozdelenie

$$V(\widehat{V(x)}) = \frac{2\sigma^4}{N^2} (N-1) \approx \frac{2\sigma^4}{N}$$

Ak použijeme Besselovu korekciu dostaneme pre ľubovoľné N

$$V(\widehat{V(x)}) = \frac{2\sigma^4}{N-1}$$

4.1.2.3 Odhad σ

$$\hat{\sigma} = \sqrt{\widehat{V(x)}} \text{ alebo jednoduchšie } \hat{\sigma} = \widehat{\sigma^2}$$

z predošlého dostaneme $\hat{\sigma} = s$

Zákon veľkých čísel garantuje konzistenciu

Rozptyl môžeme nájsť z predošlých vzťahov

$$V(\hat{\sigma}^2) = \left(\frac{d\sigma^2}{d\sigma} \right)^2 V(\sigma) = 4\sigma^2 V(\hat{\sigma})$$

ak N je primerane veľké

$$V(\hat{\sigma}) = \frac{\langle (x - \mu)^4 \rangle - \langle (x - \mu)^2 \rangle^2}{4N\sigma^2}$$

pre Gaussián sa redukuje na

$$V(\hat{\sigma}) = \frac{\sigma^2}{2N}$$

čo môžeme napísať tiež ako

$$\sigma_\sigma = \frac{\sigma}{\sqrt{2N}}$$

ak použijeme nevychýlený estimátor s

$$\sigma_\sigma = \frac{\sigma}{\sqrt{2(N-1)}}$$

Experiment - pri vhodnom usporiadaní - výsledné neistoty menšie než náhodné neistoty pre jednotlivé merania - ak zredukujeme náhodné efekty veľkým počtom meraní - predtým zanedbateľné systematické efekty sa stávajú dominantné.

4.1.2.4 Odhad korelačného koeficientu

korelačný koeficient pre vzorku - odhad korelácie zdrojového rozdelenia

$$\hat{\rho} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}$$

Besselova korekcia - odstránenie vychýlenosti

$$\hat{\rho} = r = \frac{\overline{xy} - \bar{x}\bar{y}}{(N-1)s_x s_y}$$

keď rozmer N vzorky dát je veľmi veľký neistota je potom $\sigma_\rho = \frac{1-\rho^2}{\sqrt{N-1}}$

lepšia aproximácia, ktorá pracuje aj pre menšie N je transformácia na premennú z

$$z = \frac{1}{2} \ln \frac{1+\hat{\rho}}{1-\hat{\rho}}$$

ak x a y majú Gaussovské rozdelenie potom z má viac Gaussovský tvar ako $\hat{\rho}$.

Štandardná odchýlka pre z je $1/\sqrt{N-3}$

Príklad

15 študentov fyziky písalo písomnú prácu, ktorá bola ohodnotená známkou. Korelácia medzi známkou z písomnej práce a jeho výslednou známkou zo skúšky z fyziky bola nájdená ako $r = -0.11$. Ak táto korelácia je skutočne záporná, z toho vyplýva, že literárne nadaní študenti sú slabí z fyziky a naopak. Je nejaké oprávnenie na takéto tvrdenie?

Riešenie

Transformujeme r na z a dostaneme

$$z = \frac{1}{2} \ln \left(\frac{0.89}{1.11} \right) = -0.1104$$

Neistota je $1/\sqrt{12} = 0.2887$. Takže odchýlka od nuly je menšia ako polovica štandardnej odchýlky, takže korelácia je nevýznamná.

4.2 MAXIMÁLNA VIEROHODNOSŤ (LIKELIHOOD)

Princíp maximálnej vierohodnosti (ML) je metóda pre odhad. Pre vzorku dát $\{x_1, \dots, x_N\}$ estimátor maximálnej vierohodnosti \hat{a} je taká hodnota a pre ktorú je vierohodnosť

$$L(x_1, x_2, \dots, x_N; a) = \prod P(x_i, a) \text{ maximálna.}$$

Určíme takú hodnotu a, ktorá robí pravdepodobnosť získania aktuálnych výsledkov $\{x_1, \dots, x_N\}$ takou veľkou ako je to možné. V praxi je ľahšie maximalizovať logaritmus L, ktorý je sumou potom logaritmov pravdepodobností

Príklad

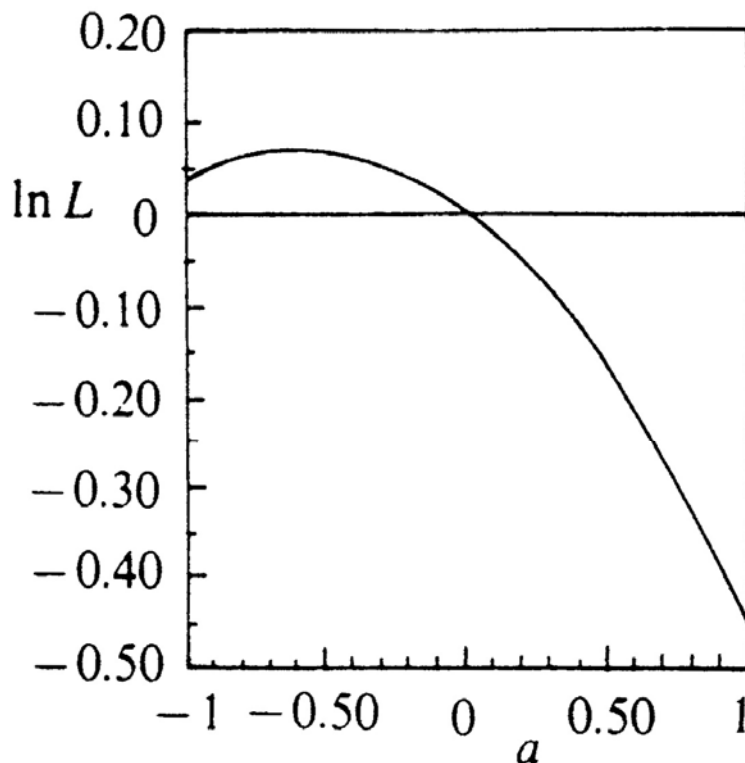
Predpokladajme, že päť hodnôt x : 0.89, 0.03, 0.50, 0.36, 0.49 bolo generovaných rozdelením, ktoré je známe a má tvar $P(x) = 1.0 + a(x - 0.5)$ medzi 0 a 1. Z týchto piatich hodnôt vypočítajte logaritmus vierohodnosti rôznych hodnôt a a zobrazíte výsledok.

Riešenie

ak $a = +1.0$ je to $\ln 1.39 + \ln 0.53 + \ln 1.0 + \ln 0.86 + \ln 0.99 = -0.47$

ak $a = -1.0$ je to $\ln 0.61 + \ln 1.47 + \ln 1.0 + \ln 1.14 + \ln 1.01 = +0.03$

Výsledky sú zobrazené na obr. 4.1



Obr.4.1 Funkcia vierohodnosti

Na obr. 4.1 je ukázaný logaritmus vierohodnosti pre hodnoty a medzi -1.0 a 1.0 . Môžeme vidieť, že maximálna vierohodnosť je pri $a = -0.6$ - to je maximálna vierohodnosť odhadu a

V mnohých prípadoch sa maximum nedá nájsť priamo, z grafu alebo podobne. Miesto toho je ľahšie nájsť polohu maxima derivovaním rovnice pre L a položením $= 0$.

$$\left. \frac{d \ln L}{da} \right|_{a=\hat{a}} = 0$$

riešime analyticky, alebo ak sa to nedá - numericky

napr. predpokladajme, že študujeme rozpad nejakého stavu s neznámou dobou života, t.j. taký stav sa rozpadá ako $(1/\tau)e^{-t/\tau}$ a chceme nájsť τ . Nameriame súbor N pozorovaných hodnôt dôb života $\{t_i\}$. Logaritmus funkcie vierohodnosti je

$$\ln L = \sum \ln \left(\frac{1}{\tau} e^{-\frac{t_i}{\tau}} \right) = \sum \left(-\frac{t_i}{\tau} - \frac{1}{\tau} \right)$$

derivovaním podľa τ dostaneme $\frac{d \ln L}{d\tau} = \sum \left(\frac{t_i}{\tau^2} - \frac{1}{\tau} \right)$

položením rovné nule dostaneme maximálne vierohodný estimátor $\hat{\tau}$

$$\sum \left(\frac{t_i}{\tau^2} - \frac{1}{\tau} \right) = 0$$

$$\hat{\tau} = \frac{1}{N} \sum t_i$$

napr. pre ukávanie flexibility metódy maximálnej vierohodnosti predpokladajme podobný experiment ako predošlý prípad s tým, že zariadenie neumožňuje meranie dlhšie než nejaký čas T . Funkcia pravdepodobnosti potom je $\frac{1}{\tau} e^{-t/\tau} (1 - e^{-T/\tau})^{-1}$.

Derivovaním logaritmu vierohodnosti a položením rovné nule dostaneme

$$\hat{\tau} = \frac{1}{N} \sum t_i + \frac{1}{N} \sum \frac{T e^{-T/\hat{\tau}}}{(1 - e^{-T/\hat{\tau}})} \quad \text{čo je potrebné riešiť numericky}$$

napr. ML a Gaussovský vážený priemer. Predpokladajme $\{x_i\}$ sú merania rovnakej veličiny ale s rôznou presnosťou, t.j. x_i sú z Gaussovskej strednej hodnoty μ (ktorú chceme nájsť) a štandardné odchýlky σ_i (sú známe - ide o neistoty meraní). Funkcia pravdepodobnosti pre x_i je

$$P(x_i; \mu, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-(x_i - \mu)^2 / 2\sigma_i^2}, \quad \text{takže logaritmus}$$

vierohodnosti potom je $\ln L = \sum_i -\ln \sigma_i \sqrt{2\pi} - \sum \frac{(x_i - \mu)^2}{2\sigma_i^2}$. Derivovaním podľa μ

$$\text{nájdem maximum} \quad \left(\frac{x_i - \hat{\mu}}{\sigma_i^2} \right) = 0 \quad \hat{\mu} = \frac{\sum x_i / \sigma_i^2}{\sum 1 / \sigma_i^2}$$

napr. ML odhad Gaussovskej strednej hodnoty μ a σ - predpokladajme súbor meraní veličiny μ keď σ je rovnaké pre všetky x_i , ale je neznáme.. In L dostaneme derivovaním podľa μ a σ takže dostaneme

$$\sum_{i=1}^n (x_i - \hat{\mu}) = 0$$

$$\sum \frac{(x_i - \hat{\mu})^2}{\hat{\sigma}^2} - \sum \frac{1}{\hat{\sigma}^2} = 0$$

Z prvej rovnice dostaneme $\hat{\mu} = \bar{x}$

$$\text{Druhá dáva} \quad \hat{\sigma}^2 = \frac{1}{N} \sum (x_i - \bar{x})^2$$

4.2.1 ML : Konzistentnosť, vychýlenosť a invariancia

ML estimátory sú obvykle (nie vždy) konzistentné.. Vo všeobecnosti sú tiež vychýlené. Vychýlenie sa významne znižuje pokiaľ rozmer vzorky je primerane veľký. Vychýlenosť je cenou za iné, výhodné vlastnosti ML metódy, najmä invariáciu voči transformáciám parametrov.

napr. pre rozptyl dostaneme, že **estimátor (σ^2) a (estimátor σ)² sú rovnaké**. To je vždy splnené pre ML estimátory. Vo všeobecnosti platí

$$\widehat{f(a)} = f(\hat{a})$$

napr. predpokladajme, že \hat{a} je estimátor, ktorý je nevychýlený, a pre ilustráciu, má rozdelenie pravdepodobnosti ktoré je symetrické okolo skutočnej hodnoty a_0 . Potom šanca byť viac než 10% príliš nízko a byť viac než 10 % príliš vysoko je rovná:

ak \hat{a} je 10% príliš vysoko, potom $\hat{a} = 1.1a_0$ a $\hat{a}^2 = 1.21a_0^2$

ak \hat{a} je 10% príliš nízko, potom $\hat{a} = 0.9a_0$ a $\hat{a}^2 = 0.81a_0^2$

teda pre premennú \hat{a}^2 je šanca byť viac než 21% príliš vysoko, a viac než 19% príliš nízko rovnaká. - rozdelenie pravdepodobnosti v \hat{a}^2 je nesymetrické a vychýlené.

4.2.2 Maximálna vierohodnosť pri veľkých N

pre $N \rightarrow \infty$ je každý konzistentný estimátor nevychýlený, takže aj vychýlenosť konzistentných ML estimátorov zaniká pri veľkých N alebo asymptotickej limite. Tieto estimátory sú tiež efektívne. Pre rozptyl \hat{a} dostaneme

$$\sigma_{\hat{a}}^2 = V(\hat{a}) = -\frac{1}{\left(\frac{d^2 \ln L}{da^2} \Big|_{a_0}\right)} = -\frac{1}{\left\langle \frac{d^2 \ln L}{da^2} \right\rangle}$$

čo je minimum hraníc rozptylu

4.2.3 Neistoty v ML estimátoroch

Pre veľké vzorky rozptyl ML estimátora je rovný MVB (**hranica minimálneho rozptylu** - minimum variance bound). To znamená, že Schwartzova nerovnosť

$\int u^2 dX \int v^2 dX \geq \left(\int uv dX\right)^2$ musí byť saturovaná (prechádza na rovnosť) a to sa stane

len keď veličiny u a v sú si priamo úmerné (pre všetky x) $v(x; a) = A(a)u(x; a)$

Takže dostaneme $\frac{d \ln L}{da} = A(a)(\hat{a} - a)$

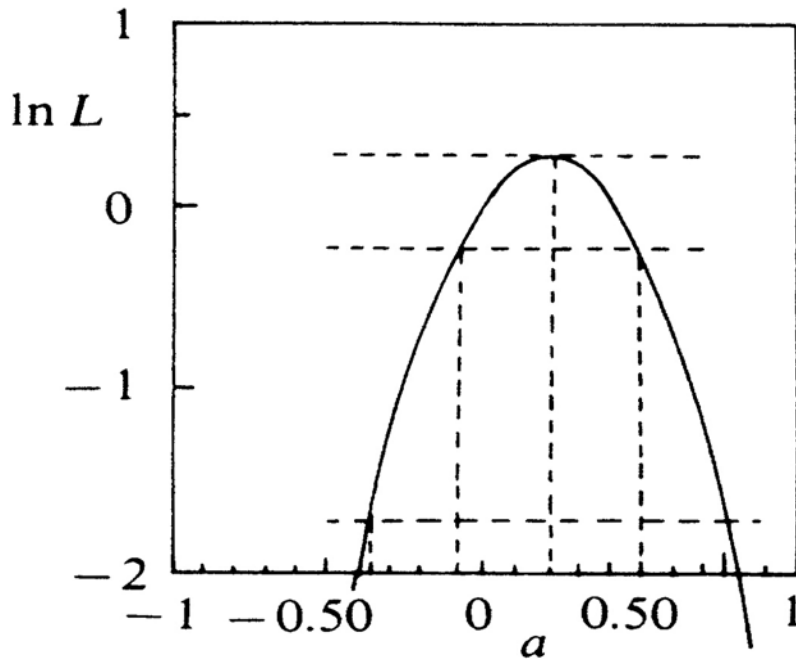
Po derivovaní podľa a a úpravách dostaneme $\left\langle \frac{d^2 \ln L}{da^2} \right\rangle = \frac{d^2 \ln L}{da^2} \Big|_{a=\hat{a}}$

čo je užitočné pre ohodnotenie $V(\hat{a})$, môže sa použiť nielen pri veľkých N ale pre každý nevychýlený, efektívny ML estimátor

Vďaka CLT rozdelenie pravdepodobnosti pre \hat{a} bolo Gaussovské. Aby to bolo skutočne pravdivé, druhá derivácia ($d^2 \ln L / da^2$) musí byť konštantná - aby to bolo - dobrou aproximáciou je ak zmeny sú veľmi malé v zodpovedajúcom intervale hodnôt a v blízkosti a_0 .

S takouto aproximáciou veličina A je konštantná takže po integrovaní a odstránení

logaritmu dostaneme pre L $L(x_1, \dots, x_N; a) \propto e^{-\frac{A(a-\hat{a}(x_1, \dots, x_N))^2}{2}}$, čo ukazuje, že limita funkcie vierohodnosti získaná z dát je Gaussián, logaritmus vierohodnosti je parabola.



Obr. 4.2 Graf logaritmu funkcie vierohodnosti ukazujúci hranice 1σ a 2σ

Štandardná odchýlka Gaussiána je práve $1/\sqrt{A}$. To je tiež štandardná odchýlka estimátora \hat{a} . Z obr. 4.2 vidieť, že pri 1σ od píku log vierohodnosti padá na hodnotu zníženú o 0.5 voči hodnote maxima, pri 2σ sa znižuje o 2, pri 3σ o 4.5, atď. Ak N nie je dostatočne veľké, funkcia vierohodnosti nie je Gauss a log vierohodnosti nie je parabola. Také prípady môžu byť ohodnotené využitím vlastnosti invariance. Hoci $\ln L(a)$ môže byť neparabolické - dá sa nájsť parameter a' pre transformáciu na parabolu. Najdeme hodnotu a' pri ktorej $\ln L$ je o 0.5 nižšie ako maximum čo označuje $\pm\sigma$ limit. Potom môžeme nájsť zodpovedajúce hodnoty a .

Podľa invariance sú aj hodnotami a pri ktorých $\ln L$ padá o 0.5 pod pík. Takže na a' môžeme zabudnúť - pre konečné N aj pre veľké N . Rozdiel - tieto hodnoty nemusia byť symetrické okolo píku hodnoty - dávajú asymetrické neistoty písané obvykle v tvare $a = 0.21_{-0.27}^{+0.29}$. 2σ limity najdeme podobne z bodov kde $\ln L$ padá o 2 oproti maximu. Treba si uvedomiť, pre konečné N , že toto nebude dvojnásobok limitu jednej štandardnej odchýlky

4.2.4 Viaceré premenné

Ak chceme ohodnotiť najlepšiu hodnotu pre celý súbor parametrov (a_1, a_2, \dots, a_n) - treba zovšeobecnenie. Nájdenie L maximálne - riešenie n rovníc

$$\frac{\partial \ln L(x_1, \dots, x_N; a_1, \dots, a_n)}{\partial a_j} = 0 \quad \text{pre všetky } j = 1, \dots, n$$

rozptyl zovšeobecníme podobne. Pre veľké N je vierohodnosť mnohorozmerný Gaussián, pre jednu premennú, rozptyl je mínus inverzia očakávanej hodnoty druhej

derivácie, pre viaceré premenné kovariančná matica je mínus inverzia matice druhých derivácií:

$$\text{cov}^{-1}(a_i, a_j) = - \left\langle \frac{\partial^2 \ln L}{\partial a_i \partial a_j} \right\rangle = \left\langle \frac{\partial \ln L}{\partial a_i} \frac{\partial \ln L}{\partial a_j} \right\rangle = - \left. \frac{\partial^2 \ln L}{\partial a_i \partial a_j} \right|_{a=\hat{a}}$$

Zobrazenie funkcie vierohodnosti je problém. Je potrebné zobrazit' dve premenné. Výhodným je použiť zobrazenie pomocou vrstevníc. Pre veľké N je funkcia dvojrozmerný Gauss, Pre malé N - rôzny tvar - možnosť viacerých lokálnych maxím. Tak ako v jednorozmernom prípade - chybu určíme najdením bodov, kde lnL je o 0.5 menšie ako maximálna hodnota - body tvoria uzavretú krivku - v zlom prípade krivky. Pre veľké N je to elipsa - 1σ limity možno určiť priemetmi elipsy na zodpovedajúce osi.

napr. rozptyl odhadu strednej hodnoty μ a σ pre Gauss

$$\ln L = \sum \frac{(x_i - \mu)^2}{2\sigma^2} - N \ln \sigma - \frac{N}{2} \ln 2\pi$$

$$\mu = \bar{x}, \quad \sigma^2 = \overline{x^2} - \mu^2$$

to potom dáva pre rozptyly:

$$\left\langle \frac{\partial^2 \ln L}{\partial \mu^2} \right\rangle = -\frac{N}{\sigma^2} \quad \left\langle \frac{\partial^2 \ln L}{\partial \sigma^2} \right\rangle = -\frac{2N}{\sigma^2} \quad \left\langle \frac{\partial^2 \ln L}{\partial \mu \partial \sigma} \right\rangle = \frac{2 \sum \langle x_i - \mu \rangle}{\sigma^3} = 0$$

pretože matica je len diagonálna, inverzia je triviálna, kovariancia medzi μ a σ je nulová, rozptyly sú

$$V(\mu) = - \left\langle \frac{\partial^2 \ln L}{\partial \mu^2} \right\rangle^{-1} = \frac{\sigma^2}{N} \quad V(\sigma) = - \left\langle \frac{\partial^2 \ln L}{\partial \sigma^2} \right\rangle^{-1} = \frac{\sigma^2}{2N}$$

Pre veľké vzorky \hat{a} má rozdelenie pravdepodobnosti ktoré je nevychýlené a normálne rozdelené okolo skutočnej hodnoty a s rozptylom rovným hranici minimálneho rozptylu.

Pre limitu veľkých N (asymptotická limita) - maximálna vierohodnosť môže dávať najlepší estimátor - pre malé N je to otázne.

Výhoda - žiadne straty informácie pri binovaní - je použitá celá experimentálna informácia.

Jednoduché ohodnotenie neistoty - bod 1σ je tam, kde lnL klesne o 0.5 oproti maximu.

Pre malé N estimátory sú obvykle vychýlené - treba vedieť tvar funkcie zdrojového rozdelenia.

4.2.5 Rozšírená maximálna vierohodnosť

v štandardnej metóde hustota pravdepodobnosti x $P(x;a)$ je normovaná na 1

$$\int P(x;a) dx = 1$$

V rozšírenej metóde táto požiadavka nie je striktná. Miesto funkcie $P(x;a)$ sa použije funkcia $Q(x;a)$ pre ktorú normalizácia nie je fixná.

$$\int Q(x; a) dx = v, \quad \text{kde } v \text{ je o\u010d\u00e1van\u00fd po\u010det eventov}$$

Je to v\u00fdhodn\u00e9 pre experimenty kde nie je zn\u00e1my po\u010det eventov - \u00e1ta bran\u00e9 ur\u010dit\u00fd \u010das - n\u00e1hodn\u00e9. V\u00fdsledn\u00e1 hodnota v zahr\u00f1uj\u00fa inform\u00e1ciu z tvaru je lep\u0161\u00edm odhadom skuto\u010dnej celkovej po\u010detnosti ako z\u00e1kladn\u00fd po\u010det pozorovaných eventov N .

Inform\u00e1cia m\u00f4\u017ee by\u0161 zahrnut\u00e1 kombin\u00e1ciou \u0161tandardnej maxim\u00e1lnej vierohodnosti so znalos\u0161\u0161ou toho \u017ee pr\u00edslu\u0161n\u00e1 $Q(x; a)$ predpoved\u00e1 v eventov v pozorovanom intervale a primerane n\u00e1sob\u00ed vierohodnos\u0161 danej datovej vzorky N eventov Poissonovou

pravdepodobnos\u0161ou z\u00edskania N eventov z priemern\u00e9ho po\u010du $v \frac{v^N}{N!} e^{-v}$. Logaritmus vierohodnosti je potom ($n!$ nezávis\u00ed na a tak\u017ee ho neuva\u017dujeme)

$$\begin{aligned} \ln L &= \sum \ln P(x_i; a) - v + N \ln v = \\ &= \sum \ln [v P(x_i; a)] - v = \\ &= \sum \ln Q(x_i; a) - v \end{aligned}$$

Vzrastom normaliz\u00e1cie Q sa zvy\u0161uje vierohodnos\u0161 kv\u00f4li rastu sumy (to rob\u00ed pozorované eventy pravdepodobnej\u0161\u00edmi) ale s\u00fačasne to zni\u017euje kv\u00f4li v (rob\u00ed to menej pravdepodobn\u00fdm, \u017ee \u017eadne in\u00e9 eventy nebud\u00fa pozorované) - pri maximaliz\u00e1cii treba naj\u0161\u0161 rovnov\u00e1hu medzi dvomi efektami. Od \u0161tandardnej vierohodnosti sa to l\u00ed\u0161i dvomi sp\u00f4sobmi - za prv\u00e9 - je tu zahrnut\u00e9 v , za druhé - faktom, \u017ee $Q(x; a)$ je normované na v , miesto na 1.

4.3 MET\u00d3DA MOMENTOV

Jedn\u00e1 sa o in\u00fa met\u00f3da odhadu.

Dan\u00e1 je funkcia rozdelenia $P(x; a)$, kde a chceme odhadn\u00fa\u0161, o\u010d\u00e1van\u00e1 hodnota priemeru

$$\langle x \rangle = \int x P(x; a) dx \quad \text{je zn\u00e1mou funkciou } a. \text{ a m\u00f4\u017ee\u0161e ohodnoti\u0161 keď d\u00e1me}$$

do rovnosti o\u010d\u00e1van\u00fa hodnotu priemeru s aktu\u00e1lnou pozorovanou hodnotou z\u00edskanou z \u00e1t

$$\langle x \rangle = \int x P(x; \hat{a}) dx = \bar{x}$$

ak m\u00e1me n nezn\u00e1mych parametrov a_1, a_2, \dots, a_n - rie\u0161ime s\u00fastavu n rovn\u00edc

$$\langle x^r \rangle = \int x^r P(x; \hat{a}_1, \dots, \hat{a}_n) dx = \bar{x}^r \quad \text{pre } r = 1, \dots, n$$

normaliza\u010dn\u00e1 podmienka

$$\int P(x; \hat{a}_1, \dots, \hat{a}_n) dx = 1$$

m\u00f4\u017ee by\u0161 pokladan\u00e1 za rovnicu pre nulov\u00fd moment

4.4 MAXIM\u00c1LNA VIEROHODNOS\u0161 A NAJMEN\u0161IE \u0160TVORCE

Predpokladajme vzorku \u00e1t - s\u00fabor p\u00e1rov (x, y) - $\{(x_i, y_i)\}$ - hodnoty x s\u00fa zn\u00e1me presne, hodnoty y s\u00fa meran\u00e9 s neistotou σ_i , y m\u00f4\u017ee\u0161e vyjadri\u0161 ako funkciu $f(x)$ - z\u00e1vis\u00ed na parametri a ktor\u00fd chceme ohodnoti\u0161

Z centr\u00e1lnej limitnej vety m\u00f4\u017ee\u0161e uk\u00e1za\u0161, \u017ee y hodnoty s\u00fa Gaussovsky rozdelen\u00e9 okolo ide\u00e1lnej hodnoty, pravdepodobnos\u0161 ur\u010dit\u00e9ho y_i pre dan\u00e9 x_i je

$$P(y_i; a) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-[y_i - f(x_i; a)]^2 / 2\sigma_i^2}$$

Logaritmus vierohodnosti pre celkový súbor dát je

$$\ln L = \frac{1}{2} \sum \left[\frac{y_i - f(x_i)}{\sigma_i} \right]^2 - \sum \ln \sigma_i \sqrt{2\pi}$$

Maximalizácia vierohodnosti vedie k minimalizácii veličiny

$$\sum \left[\frac{y_i - f(x_i)}{\sigma_i} \right]^2$$

t.j. robeniu sumy štvorcov vážených rozdielov takou malou ako je možné a to je metóda najmenších štvorcov

4.5 STRATIFIKOVANÉ VZORKOVANIE

Predpokladajme, že chceme nájsť priemer nejakej veličiny pre populáciu na základe relatívne malej vzorky - napr. priemernú výšku všetkých študentov navštevujúcich určitú univerzitu - urobíme N meraní - určíme strednú výšku μ - ohodnotíme vzorkovú štandardnú odchýlku s - výsledok dostaneme v tvare $\mu \pm s/\sqrt{N}$

Môžeme to urobiť lepšie

Študenti - 2 typy - muži a ženy - je známe, že priemerné výšky mužov a žien sú rozdielne - určíme relatívne zastúpenie mužov a žien medzi študentami (Manchester - 64% : 36%).

Pomer vo vzorke fluktuuje okolo nejakej hodnoty - tieto fluktuácie prispievajú k fluktuáciám priemernej výšky - zbaviť sa ich môžeme meraním vzorky obsahujúcej určitý pomer muži : ženy - tak redukuje nepresnosť výsledku - toto sa nazýva **stratifikované vzorkovanie**

Pre kvantifikáciu - pozrime čo sa stane pri náhodnom vzorkovaní

Predpokladajme - rozdelenie je zmesou $P_1(x)$ a $P_2(x)$ s μ_1 a μ_2 v pomere f_1 a f_2 ($f_1 + f_2 = 1$) a celkový priemer $\mu = f_1\mu_1 + f_2\mu_2$. Ak urobíme N meraní očakávaný rozptyl celkového priemeru je

$$V(x) = \int [f_1 P_1(x) + f_2 P_2(x)] (x - \mu)^2 dx$$

to je práve f_1 krát očakávaná hodnota $(x - \mu)^2$ pre typ 1 a to isté pre typ 2. Potom dostaneme

$$V(x) = f_1 V_1(x) + f_2 V_2(x) + f_1 f_2 (\mu_1 - \mu_2)^2$$

ak použijeme stratifikované vzorkovanie v pomere f_1 k f_2 , CLT dáva rozptyl ako ukazujú prvé dva členy bez tretieho, tretí člen je kvôli možným (binomickým) fluktuáciám vo vzorkovaní. Stratifikované vzorkovanie to eliminuje

Vo všeobecnosti, ak vieme, že populácia je rozdelená do dvoch (alebo viacerých) typov u ktorých vieme proporcie a tieto typy majú rozdielne stredné hodnoty - je výhodné vzorkovať fixný počet z každého z týchto typov. To znižuje rozptyl vo vzorke spôsobený možným výberom rôzneho počtu z každého typu.

Toto funguje ak vieme proporciálne zastúpenie typov a ak rozdiely medzi strednými hodnotami rôznych typov sú významné.

Čo je najlepšia stratifikácia?

Predpokladajme m_1 meraní typu 1 a m_2 meraní typu 2.
Odhadneme μ_1 a μ_2 s rozptylmi V_1/m_1 a V_2/m_2 . Odhady dajú

$$\hat{\mu} = f_1 \hat{\mu}_1 + f_2 \hat{\mu}_2$$

rozptyl je

$$V = \frac{f_1^2 V_1}{m_1} + \frac{f_2^2 V_2}{m_2},$$

chceme vybrať m_1 a m_2 tak, aby rozptyl bol minimálny - $m_1 + m_2$ je obmedzené fixným N , $m_2 = N - m_1$, derivujeme rozptyl podľa m_1

$$\frac{m_1}{m_2} = \frac{f_1 \sqrt{V_1}}{f_2 \sqrt{V_2}} = \frac{f_1 \sigma_1}{f_2 \sigma_2}$$

Ak σ je rovnaké, výsledok nám potvrdzuje, že vzorky je najlepšie vybrať v pomere v akom sú dáta

KONTROLNÉ OTÁZKY

39. Čo sú to estimátory?
40. Aké sú základné vlastnosti dobrých estimátorov?
41. Ako vyjadríme funkciu vierohodnosti a čo vyjadruje?
42. Aké sú hranice minimálneho rozptylu estimátora?
43. Uvedte niektoré základné estimátory.
44. Čo vyjadruje maximálna vierohodnosť a na čo sa využíva?
45. Ako by ste určili hranice 1σ rozptylu funkcie $\ln L$?
46. Č je to metóda momentov?
47. Na čom je založené stratifikované vzorkovanie a v čom je výhodnosť jeho použitia?

ÚLOHY

4.1 Daných je 7 čísel:

20.0 19.7 20.6 18.5 21.2 20.8 20.7

a) vypočítajte priemer. Aká je neistota priemeru, ak čísla reprezentujú merania jednej veličiny s rozlíšením 0.8?

b) Odhadnite štandardnú odchýlku ak skutočná stredná hodnota je 20.0. Aká je neistota odhadu?

c) Odhadnite štandardnú odchýlku ak nevieme skutočnú strednú hodnotu.

d) Najdite neistotu priemeru ak nevieme štandardnú odchýlku.

4.2 Ak máte dve nezávislé merania rovnakej presnosti, jedno $\sin \Theta$, druhé $\cos \Theta$. Najdite metódou maximálnej vierohodnosti odhad pre Θ .

4.3 Ukážte, že ak doba života τ v exponenciálnom rozpade je najdená metódou maximálnej vierohodnosti, rozptyl pre veľké vzorky je $V(\tau) = \frac{\tau^2}{N}$

SÚHRN

Estimátor - procedúra ktorú keď aplikujeme na vzorku dát dáva číselnú hodnotu pre charakteristiku zdrojového rozdelenia Dobrý estimátor musí byť : - **konzistentný** , - **nevychýlený** , - **účinný**

predpokladajme, že veličina ktorú meriame je a , jej estimátor označujeme strieškou nad príslušnú veličinu \hat{a} .

Estimátor je **konzistentný** ak smeruje k skutočnej hodnote ak dáta idú do nekonečna

$$\lim_{N \rightarrow \infty} \hat{a} = a$$

Estimátor je **nevychýlený** ak jeho očakávaná hodnota je rovná skutočnej hodnote $\langle \hat{a} \rangle = a$

Estimátor je **účinný**, ak jeho rozptyl je malý

Pravdepodobnosť určitého súboru dát $\{x_1, x_2, x_3, \dots, x_N\}$ je súčin individuálnych pravdepodobností - tento súčin sa volá **vierohodnosť** $L(x_1, x_2, x_3, \dots, x_N; a)$ je kombinovanou pravdepodobnosťou alebo hustotou pravdepodobnosti ktorá tento určitý súbor x_i bude vytvárať z tejto hodnoty a

$$L(x_1, x_2, x_3, \dots, x_N; a) = P(x_1; a)P(x_2; a) \dots P(x_N; a) = \prod P(x_i; a)$$

Očakávaná hodnota pre každú funkciu $z \{x_i\}$ sa nájde integrovaním cez všetky možné hodnoty všetkých x_i , váhovaním celkovou pravdepodobnosťou, t.j vierohodnosťou

$$\langle f(x_1, x_2, \dots, x_N) \rangle = \int \dots \int f(x_1, \dots, x_N) L(x_1, x_2, \dots, x_N; a) dx_1 \dots dx_N$$

v obvyklom označovaní $\langle f(x_1, \dots, x_N) \rangle = \int f L dX$

platí $\langle \hat{a} \rangle = \int \hat{a} L dX$ $\langle \hat{a}^2 \rangle = \int \hat{a}^2 L dX$

požiadavka konzistencie je $\lim_{N \rightarrow \infty} \langle \hat{a} - a \rangle = 0$

hranica minimálneho rozptylu (MVB- minimum variance bound) pre nevychýlený

estimátor je $V(\hat{a}) \geq \frac{1}{\left\langle \left(\frac{d \ln L}{da} \right)^2 \right\rangle}$ alebo inak $V(\hat{a}) \geq \frac{-1}{\left\langle \left(\frac{d^2 \ln L}{da^2} \right) \right\rangle}$

ak pre nejaký estimátor - $V(\hat{a})$ je rovné MVB potom je estimátor \hat{a} účinný, ak nie - jeho účinnosť je $MVB/V(\hat{a})$

Princíp maximálnej vierohodnosti (ML) je metóda pre odhad. Pre vzorku dát $\{x_1, \dots, x_N\}$ estimátor maximálnej vierohodnosti \hat{a} je taká hodnota a pre ktorú je vierohodnosť $L(x_1, x_2, \dots, x_N; a) = \prod P(x_i, a)$ maximálna. Určíme takú hodnotu a , ktorá robí pravdepodobnosť získania aktuálnych výsledkov $\{x_1, \dots, x_N\}$ takou veľkou ako je to možné. V praxi je ľahšie maximalizovať logaritmus L , ktorý je sumou potom logaritmov pravdepodobností

pre $N \rightarrow \infty$ je každý konzistentný estimátor nevychýlený, takže aj vychýlenosť konzistených ML estimátorov zaniká pri veľkých N alebo asymptotickej limite. Tieto estimátory sú tiež efektívne.

Jednoduché ohodnotenie neistoty - hranice 1σ sú tam, kde $\ln L$ klesne o $1/2$ oproti maximu.

Metóda momentov je ďalšia metóda odhadu. Ak máme n neznámych parametrov a_1, a_2, \dots, a_n - riešime sústavu n rovníc

$$\langle x^r \rangle = \int x^r P(x; \hat{a}_1, \dots, \hat{a}_n) dx = \overline{x^r} \quad \text{pre } r = 1, \dots, n$$

normalizačná podmienka

$$\int P(x; \hat{a}_1, \dots, \hat{a}_n) dx = 1$$

môže byť pokladaná za rovnicu pre nulový moment

Stratifikované vzorkovanie - ak vieme, že populácia je rozdelená do dvoch (alebo viacerých) typov u ktorých vieme proporcie a tieto typy majú rozdielne stredné hodnoty - je výhodné vzorkovať fixný počet z každého z týchto typov. To znižuje rozptyl vo vzorke spôsobený možným výberom rôzneho počtu z každého typu. Toto funguje ak vieme proporciálne zastúpenie typov a ak rozdiely medzi strednými hodnotami rôznych typov sú významné.

5 METÓDA NAJMENŠÍCH ŠTVORCOV

UČEBNÉ CIELE

Cieľom tejto kapitoly je bližšie oboznámiť študentov s najrozšírenejšou metódou fitovania dát, naučiť optimálne používať túto metódu pre rôzne funkčné závislosti, vysvetliť χ^2 rozdelenie, popísať nelineárnu metódu najmenších štvorcov, problémy pri fitovaní dát.

KLÚČOVÉ SLOVÁ

metóda najmenších štvorcov, fitovanie dát, lineárny fit, χ^2 rozdelenie, nelineárna metóda najmenších štvorcov

Metóda je použiteľná - ak máme dve premenné x, y

- súbor presne známych x hodnôt

- zodpovedajúci súbor y hodnôt meraných s presnosťou σ

- funkciu $f(x; a)$ ktorá predpovedá hodnoty y pre každé x ,

chceme určiť parameter a

Princíp metódy - dá sa odvodiť z princípu maximálnej vierohodnosti - minimalizuje sa suma χ^2

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - f(x; a)}{\sigma_i} \right]^2$$

$$\frac{d\chi^2}{da} = 0$$

$$\sum_i \frac{1}{\sigma_i^2} \frac{df(x; a)}{da} [y_i - f(x; a)] = 0$$

5.1 JEDNODUCHÁ ÚMERNOSŤ

Uvažujme funkciu v tvare $y = m x$

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - f(x_i; a)}{\sigma_i} \right]^2$$

derivovaním podľa m dostaneme

$$\frac{d\chi^2}{dm} = \sum (-2x_i) \frac{y_i - mx_i}{\sigma_i^2}$$

ak sú neistoty rovnaké - dostaneme $\frac{-2}{\sigma_i^2} \sum (x_i y_i - mx_i^2)$, minimalizujeme

$$\sum (x_i y_i - \hat{m} x_i^2) = 0$$

$$\sum x_i y_i = \hat{m} \sum x_i^2, \quad \text{predelíme } N \text{ dostaneme } \hat{m} = \frac{\overline{xy}}{\overline{x^2}}$$

Ak výsledok napíšeme v tvare $\hat{m} = \sum \frac{x_i}{N x^2} y_i$, môžeme ohodnotiť presnosť

$$V(\hat{m}) = \sum_i \left(\frac{x_i}{N x^2} \right)^2 \sigma^2 = \frac{\sigma^2}{N x^2}$$

5.2 LINEÁRNY FIT

Uvažujme funkciu v tvare $y = mx + c$

Aj naďalej predpokladáme, že neistoty sú rovnaké.

Minimalizujeme sumu

$$\sum_i (y_i - mx_i - c)^2$$

derivovaním podľa c a dostaneme

$$\sum_i -2(y_i - \hat{m}x_i - \hat{c}) = 0$$

predelením N dostaneme

$$\bar{y} - \hat{m}\bar{x} - \hat{c} = 0$$

Derivovaním sumy podľa m

$$\sum_i -2x_i (y_i - \hat{m}x_i - \hat{c}) = 0$$

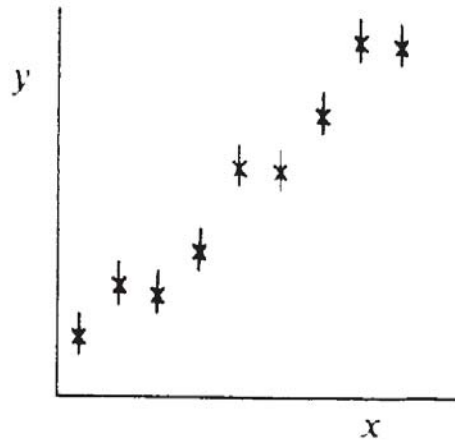
a potom

$$\overline{xy} - \hat{m}\overline{x^2} - \hat{c}\bar{x} = 0$$

vyjadrením c a dosadením dostaneme smernicu

$$\hat{m} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}$$

môžeme ju vyjadriť v pojmoch kovariancie a rozptylu



Obr. 5.1 Obvyklý problém pri fitovaní dát

$$\hat{m} = \frac{\text{cov}(x, y)}{V(x)}$$

dosadením dostaneme pre c

$$\hat{c} = \frac{\overline{x^2 y} - \bar{x} \bar{xy}}{x^2 - \bar{x}^2}$$

alebo vo výhodnejšom tvare

$$\hat{c} = \bar{y} - \hat{m} \bar{x}$$

rozptyl odhadu m dostaneme -vyjadríme si m v tvare

$$\hat{m} = \sum_i \frac{x_i - \bar{x}}{N(x^2 - \bar{x}^2)} y_i$$

z pravidiel kombinácie neistôt vyplýva ($V(f) \sim (df/dy)^2 V(y)$)

$$V(\hat{m}) = \sum_i \left[\frac{x_i - \bar{x}}{N(x^2 - \bar{x}^2)} \right]^2 \sigma^2$$

z toho

$$V(\hat{m}) = \frac{\sigma^2}{N(x^2 - \bar{x}^2)}$$

Podobne dostaneme rozptyl pre c

$$V(\hat{c}) = \sum_i \left[\frac{\bar{x}^2 - \bar{x} y_i}{N(x^2 - \bar{x}^2)} \right]^2 \sigma^2 = \frac{\sigma^2 \bar{x}^2}{N(x^2 - \bar{x}^2)}$$

pre kovarianciu platí $\text{cov}(\hat{m}, \hat{c}) = -\frac{\bar{x}}{N(x^2 - \bar{x}^2)} \sigma^2$

spätným dosadením dostaneme pre χ^2

$$\chi^2 = N \frac{V(y)}{\sigma^2} (1 - \rho_{xy}^2)$$

Veličina $V(y) = \overline{y^2} - \bar{y}^2$ popisuje rozmazanie y hodnôt celého súboru dát, zatiaľ čo σ popisuje rozmazanie jednotlivých hodnôt y okolo jej skutočnej hodnoty

5.3 VÁHOVANÝ LINEÁRNY FIT

Ak σ_i nie sú rovnaké pre všetky dáta minimalizuje sa suma $\sum_i \frac{(y_i - mx_i - c)^2}{\sigma_i^2}$

vzťahy pre m a c sú také isté ako predtým, len normalizácia nie je počtom bodov N, ale celkovou váhou $\sum 1/\sigma_i^2$

nahradíme

$$\frac{\sum y_i}{N} \rightarrow \frac{\sum y_i / \sigma_i^2}{\sum 1 / \sigma_i^2}$$

vo vzťahoch pre rozptyl nahradíme σ^2

$$\frac{\sigma^2}{\sigma^2} = \frac{\sum_i \sigma_i^2 / \sigma_i^2}{\sum_i 1 / \sigma_i^2} = \frac{N}{\sum_i 1 / \sigma_i^2}$$

Extrapolácia (interpolácia) hodnôt - máme X, predpovedané Y je $\hat{m}x + \hat{c}$, neistota je daná $V(Y) = V(\hat{c}) + XV(\hat{m}) + 2X \text{cov}(\hat{m}, \hat{c})$ - dôležitý je posledný člen

Nepresnosť extrapolácie je

$$V(Y) = \frac{\sigma^2 (X - \bar{x})^2}{N(x^2 - \bar{x}^2)} + \frac{\sigma^2}{N}$$

5.4 SYSTEMATICKÉ NEISTOTY A LINEÁRNY FIT

Uvažujme prípad, keď všetky hodnoty y majú náhodnú neistotu σ a systematickú S. sú vyjadrené uvedenými vzťahmi, úplný vzťah pre neistotu je

$$V(\hat{m}) = \frac{1}{N^2 (\bar{x}^2 - \bar{x}^2)^2} \sum_i \sum_j (x_i - \bar{x})(x_j - \bar{x}) \text{cov}(y_i, y_j)$$

$\text{cov}(y_i, y_j) = \delta_{ij} \sigma^2 + S^2$, takže

$$V(\hat{m}) = \frac{1}{N^2 (\bar{x}^2 - \bar{x}^2)^2} \left[\sum_i (x_i - \bar{x})^2 \sigma^2 + \sum_i \sum_j (x_i - \bar{x})(x_j - \bar{x}) S^2 \right]$$

Prvá suma dáva štandardné vyjadrenie

$$V(\hat{m}) = \frac{\sigma^2}{N(\bar{x}^2 - \bar{x}^2)} \text{ (už uvedené), druhé vypadnú pretože } \frac{1}{N} \sum x_i = \bar{x}, \text{ takže pridanie}$$

systematickej neistoty pre y hodnoty nemá na smernicu vplyv pre konštantu dostaneme

$$V(\hat{c}) = \frac{1}{N^2 (\bar{x}^2 - \bar{x}^2)^2} \sum_i \sum_j (\bar{x} - x_i)(\bar{x} - x_j) \text{cov}(y_i, y_j)$$

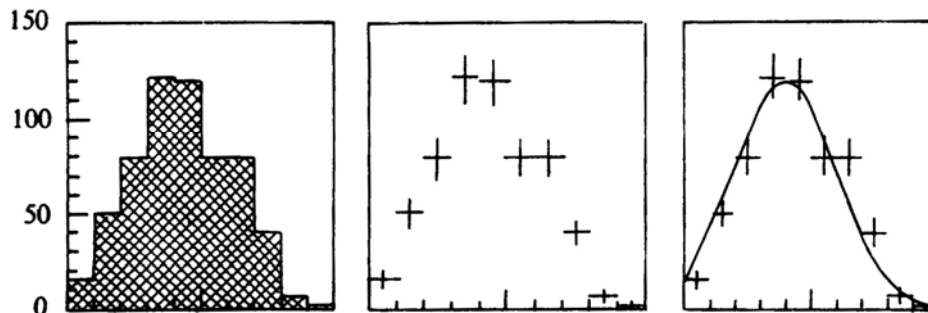
Systematická neistota prispieva na posun dodatočným členom, ktorý je rovný práve S^2

5.5 FITOVANIE BINOVANÝCH DÁT

predpokladajme N eventov, funkciu pravdepodobnosti $P(x;a)$, dáta v binoch $1, \dots, N_B$, číslo binu j - centrovane na bod x_j , šírka binu W_j , bin obsahuje n_j eventov ideálny počet eventov očakávaný v bine j je $f_j = NW_j P(x_j; a)$.

Aktuálny počet bude popísaný Poissonovou štatistikou, takže štvorec chyby je rovný priemeru a celkový χ^2 sumovaný cez všetky biny bude

$$\chi^2 = \sum_j \frac{(n_j - f_j)^2}{f_j}$$



Obr.5.2 Fitovanie histogramu

reálne - aproximácia

$$\chi^2 \approx \sum_j \frac{(n_j - f_j)^2}{n_j}$$

je ľahšie ohodnotiť numericky - aproximácia platí, len keď n_j a f_j sú veľké a nie príliš rozdielne - nezabudnúť zahrnúť šírku binu W keď pracujeme mimo f_j .

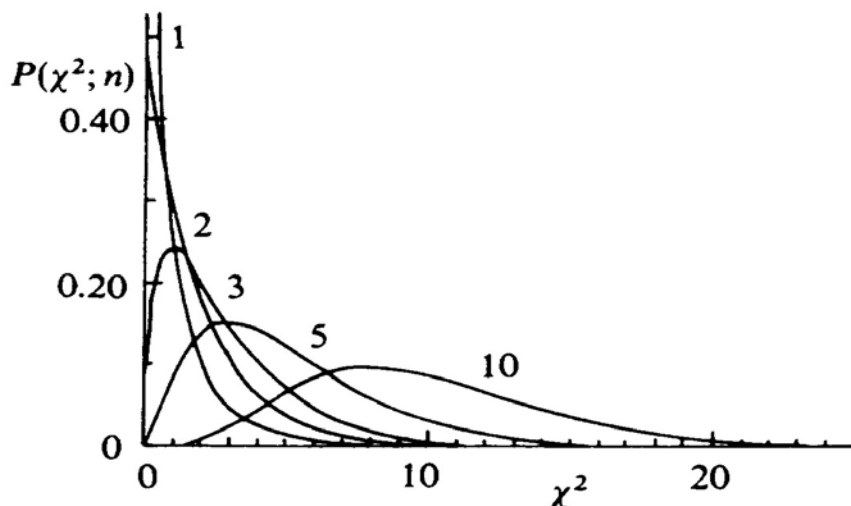
5.6 χ^2 ROZDELENIE

χ^2 je definované ako štvorec rozdielov medzi pozorovanými hodnotami a teoretickou predpoveďou, váhované chybami merania

$$\chi^2 = \sum_i \frac{[y_i - f(x_i)]^2}{\sigma_i^2} = \sum_{i=1}^N \left(\frac{y_i^{\text{aktual}} - y_i^{\text{ideal}}}{\text{očakávaná chyba}} \right)^2$$

Ak funkcia súhlasí s aktuálnou hodnotou - χ^2 je malé, veľká hodnota nie dobrá, taktiež príliš malá hodnota nie dobrá, môže znamenať, že chyby boli nadhodnotené
Rozdelenie pravdepodobnosti pre χ^2

$$P(\chi^2; n) = \frac{2^{-n/2}}{\Gamma(n/2)} \chi^{n-2} e^{-\chi^2/2}$$



Obr. 5.3 Niektoré χ^2 rozdelenia

$\Gamma(x)$ štandardná gama funkcia

$$\Gamma(1) = 1 \quad \Gamma(1/2) = \sqrt{\pi} \quad \Gamma(n+1) = n\Gamma(n)$$

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} \ln^n t dt$$

pre $n = 0, 1, 2, \dots$ $\Gamma(n+1) = n!$

pre $n = 1/2, 3/2, 5/2, \dots$ $\Gamma(n+1) = n(n-1)(n-2)\dots(3/2\sqrt{\pi})$

n je počet stupňov volnosti = N - počet premenných, ktoré treba nastaviť pre minimalizáciu χ^2 .

χ^2 rozdelenie má strednú hodnotu n a rozptyl $2n$

χ^2 na stupeň volnosti by malo byť okolo jednotky - z centrálnej limitnej vety - χ^2 rozdelenie pre veľké n ide ku Gaussu (n musí byť extrémne veľké). Lepšie použiť

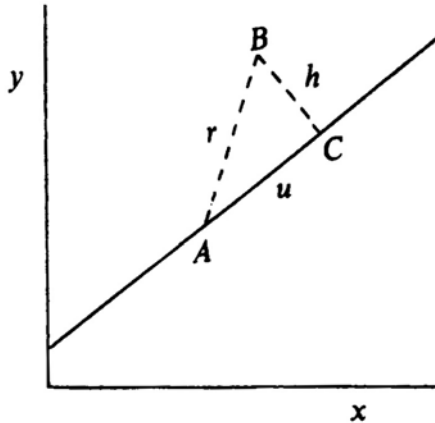
$\sqrt{2\chi^2}$ - rozdelenie je rozumnejšou aproximáciou Gaussu so strednou hodnotou $\sqrt{2n-1}$ a jednotkovým rozptylom pre hodnoty n väčšie ako ~ 30

napr: vhodnosť fitu a χ^2 . Máme frekvenčné rozdelenie, ktoré obsahuje veľký pík zahrňujúci 45 binov. Fitovanie Gaussom dáva $\chi^2 = 73$. Sú tu 3 parametre (stredná hodnota, šírka, a normalizácia na Gauss), počet stupňov volnosti je 42. $\sqrt{2n-1}$ dáva 9.1, $\sqrt{2\chi^2}$ je 12.1 čo je o 3 jednotky väčšie, čo je tiež 3σ . Z toho je vidieť, že pík nie je dobre popísaný Gaussiánom.

Môže to byť použité aj opačne. Predpokladajme súbor meraní s rovnakou presnosťou, ale neznámou. Štvorec odchýlok na stupeň volnosti dáva σ^2 , z čoho môžeme ohodnotiť chybu merania nemôžeme však už použiť χ^2 ako test fitu

5.7 CHYBY U X A J Y

Predpokladajme, že aj x aj y sú zaťažené chybou merania, pre jednoduchosť rovnakými s hodnotou σ a fitujeme priamkou



Obr. 5.4 B je bod merania. Plná čiara zodpovedá $y=f(x;a)$, čo je v tomto prípade priamka

použijeme metódu maximálnej vierohodnosti, meraný bod B na obr. môže pochádzať z ľubovoľného bodu A ideálnej priamky. Hustota pravdepodobnosti

pre takýto prípad je $P(A \rightarrow B) = \frac{1}{2\pi\sigma^2} e^{-(x_A-x_B)^2/2\sigma^2} e^{-(y_A-y_B)^2/2\sigma^2}$

čo je práve $\frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2}$ a tiež $\frac{1}{2\pi\sigma^2} e^{-u^2/2\sigma^2} e^{-h^2/2\sigma^2}$

Aby sme určili celkovú pravdepodobnosť, že B pochádza z A, musíme integrovať $P(A \rightarrow B)$ cez všetky u . Integrál $\exp(-u^2/2\sigma^2)$ dáva konštantu $\sigma\sqrt{2\pi}$, takže pravdepodobnosť pozorovania bodu v B je úmerná $\exp(-h^2/2\sigma^2)$. h je kolmá vzdialenosť z bodu ku priamke, t.j. z nameraného bodu B do bodu C, čo je bod na priamke s najväčšou pravdepodobnosťou, že z neho pochádza bod B.

Vytvoríme χ^2 pre minimalizáciu sumy štvorcov vzdialeností každého meraného bodu k bodu na priamke z ktorého najpravdepodobnejšie pochádza. To, že môže pochádzať z iného bodu nemusíme uvažovať. Ak chyby nie sú rovnaké nahradíme $y' = y/\sigma_y$ a $x' = x/\sigma_x$

pre $y = mx + c$ a rovnake chyby

$$h_i = \frac{y_i - mx_i - c}{\sqrt{1+m^2}}$$

potom $\chi^2 \propto \sum_i \frac{(y_i - mx_i - c)^2}{1+m^2}$ derivovaním podľa c dostaneme $\bar{y} = \hat{m}\bar{x} + \hat{c}$

derivovaním podľa m dostaneme výsledok $\hat{m} = A \pm \sqrt{A^2 + 1}$ kde $A = \frac{V(y) - V(x)}{2 \text{cov}(x, y)}$

dve riešenia pre smernicu sú vzájomne kolmé, jedno dáva najlepšiu priamku a druhé najhoršiu. Znamienko plus je pre kladnú kovarianciu minus pre zápornú

Ak chyby nie sú rovnaké

$$\hat{m} = \frac{\sigma_y}{\sigma_x} \left(A \pm \sqrt{A^2 + 1} \right) \quad A = \frac{\sigma_x^2 V(y) - \sigma_y^2 V(x)}{2\sigma_x \sigma_y \text{cov}(x, y)}$$

Ak sú chyby rôzne pre každý bod - niet analytického riešenia - riešenie je možné nájsť len numericky

5.8 NELINEÁRNA METÓDA NAJMENŠÍCH ŠTVORCOV

Ak funkcia $f(x; \mathbf{a})$ nie je lineárna pre \mathbf{a}_r - vo všeobecnosti sa musí použiť iteratívna technika. Urobí sa prvý odhad \mathbf{a}^0 , potom gradienty sú (pre nezávislé merania)

$$\left. \frac{\partial \chi^2}{\partial a_r} \right|_{\mathbf{a}=\mathbf{a}^0} = g_r(\mathbf{a}^0) = \sum_i -\frac{2}{\sigma_i^2} [y_i - f(x_i; \mathbf{a}^0)] \frac{\partial f(x_i; \mathbf{a}^0)}{\partial a_r}$$

chceme nájsť prírastok $\delta \mathbf{a}$ takže $g_r(\mathbf{a}^0 + \delta \mathbf{a}) = \left. \frac{\partial \chi^2}{\partial a_r} \right|_{\mathbf{a}=\mathbf{a}^0 + \delta \mathbf{a}} = 0$ pre všetky r

Rozvinieme do Taylorovho radu, pričom zachováme len členy nulového a prvého

$$\text{rádu} \quad g_r(\mathbf{a}^0 + \delta \mathbf{a}) \approx g_r(\mathbf{a}^0) + \sum_s \frac{\partial g_r}{\partial a_s} \delta a_s = g_r(\mathbf{a}^0) + \sum_s \frac{\partial^2 \chi^2}{\partial a_r \partial a_s} \delta a_s$$

keď pre zjednodušenie zapíšeme $f(x_i; \mathbf{a}^0)$ ako f_i a

$$G_{rs} = \frac{\partial^2 \chi^2}{\partial a_r \partial a_s} = \sum_i \left(-\frac{2}{\sigma_i^2} \right) \left[-\left(\frac{\partial f_i}{\partial a_r} \right) \left(\frac{\partial f_i}{\partial a_s} \right) + (y_i - f_i) \left(\frac{\partial^2 f_i}{\partial a_r \partial a_s} \right) \right]$$

$\delta \mathbf{a}$ nájdeme z maticovej rovnice $\delta \mathbf{a} = -\mathbf{G}^{-1} \mathbf{g}$

riešenie je možné nájsť iteráciou.

V iteračnom riešení - dobrý prvý odhad reprezentuje 90% procesu. Vytvoríme matice a invertujeme. Keď g_r sa stane dostatočne malé, iteráciu zastavíme a výsledné \mathbf{a} je riešením - využitie softwarových balíkov

ÚLOHY

5. 1 Vozík sa pohybuje po dráhe konštantnou rýchlosťou, ktorú potrebujeme merať. Prechádza cez bod $d = 0$ presne v čase $t = 0$. V rovnakých časových intervaloch, určených stroboskopom, fotografujeme vozík a tak môžeme zmerať jeho polohu s presnosťou 2mm. Výsledky sú:

Čas [s]: 1.0 2.0 3.0 4.0 5.0 6.0

Vzdialenosť [mm]: 11 19 33 40 49 61

Nájdite rýchlosť, neistotu takéhoto určenia a χ^2 .

5. 2 Určujeme zrýchlenie spôsobené gravitáciou vypínaním elektromagnetu držiaceho oceľovú guľičku a meraním času pádu na fixnú vzdialenosť d , vzdialenosť je meraná presne, čas s presnosťou 0.01 s.

Čas [s]: 0.16 0.40 0.58 0.72 0.97

Vzdialenosť [mm]: 0.20 1.00 2.00 3.00 5.00

Vypočítajte gravitačné zrýchlenie so zodpovedajúcou neistotou

a) ak časy sú také ako v tabulke, b) pole v elektromagnete spôsobuje konštantné oneskorenie pri vypúšťaní ocelevej guľičky. Okomentujte rozdiely a χ^2 oboch fitov.

5. 3 Ak vieme, že $y = ax + b\sin x$ najdite metódou najmenších štvorcov odhad pre a a b

x (radián) : 0.2 0.3 0.4 0.5 0.7 0.7

y : 0.599 0.896 1.189 1.479 1.756 2.044

5. 4 Rozpadajúci sa rádioaktívny zdroj je meraný pomocou GM počítača. Dáta sú merané krátky čas (1min.) v hodinových intervaloch. Určite polčas premeny.

Čas [hodiny] : 0 1 2 3 4 5 6 7 8

Počet imp. : 997 520 265 127 70 35 16 7 3

KONTROLNÉ OTÁZKY

48. V čom spočíva princíp metódy najmenších štvorcov?
49. Ako môžeme vyjadriť odhad parametrov lineárneho fitu metódou najmenších štvorcov za predpokladu, že neistoty experimentálnych dát sú rovnaké?
50. Ako vyjadríme rozptyl týchto parametrov?
51. Ako sa prejaví vplyv váhovania experimentálnych dát neistotami vo fite?
52. Ako prispieva k rozptylu parametrov systematická neistota?
53. Aké sú základné charakteristiky χ^2 rozdelenia?

SÚHRN

Metóda najmenších štvorcov je použiteľná napr.: - ak máme dve premenné x, y , - súbor presne známych x hodnôt, - zodpovedajúci súbor y hodnôt meraných s presnosťou σ , - funkciu $f(x; a)$ ktorá predpovedá hodnoty y pre každé x , parameter a chceme určiť.

Princíp metódy - dá sa odvodiť z princípu maximálnej vierohodnosti - minimalizuje sa suma χ^2

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - f(x_i; a)}{\sigma_i} \right]^2$$

$$\frac{d\chi^2}{da} = 0$$

$$\sum_i \frac{1}{\sigma_i^2} \frac{df(x; a)}{da} [y_i - f(x; a)] = 0$$

Lineárny fit $y = mx + c$, rovnaké neistoty: $\hat{m} = \frac{\overline{xy} - \bar{x}\bar{y}}{x^2 - \bar{x}^2}$, $\hat{c} = \frac{\overline{x^2 y} - \bar{x}\overline{xy}}{x^2 - \bar{x}^2}$

rozptyl $V(\hat{m}) = \frac{\sigma^2}{N(x^2 - \bar{x}^2)}$, $V(\hat{c}) = \frac{\sigma^2 \bar{x}^2}{N(x^2 - \bar{x}^2)}$

pre kovarianciu platí $\text{cov}(\hat{m}, \hat{c}) = -\frac{\bar{x}}{N(x^2 - \bar{x}^2)} \sigma^2$

spätným dosadením dostaneme pre χ^2 $\chi^2 = N \frac{V(y)}{\sigma^2} (1 - \rho_{xy}^2)$

Vážený lineárny fit - ak σ_i nie sú rovnaké pre všetky dáta minimalizuje sa suma

$$\sum_i \frac{(y_i - mx_i - c)^2}{\sigma_i^2}, \text{ vo vzťahoch pre } m \text{ a } c \text{ nahradíme } \frac{\sum y_i}{N} \rightarrow \frac{\sum y_i / \sigma_i^2}{\sum 1 / \sigma_i^2}$$

vo vzťahoch pre rozptyl nahradíme σ^2 $\sigma^2 = \frac{\sum \sigma_i^2 / \sigma_i^2}{\sum 1 / \sigma_i^2} = \frac{N}{\sum 1 / \sigma_i^2}$

Systematické neistoty - Keď hodnoty y majú náhodnú neistotu σ a systematickú S .

$$V(\hat{m}) = \frac{1}{N^2 (x^2 - \bar{x}^2)^2} \sum_i (x_i - \bar{x})^2 \sigma^2 + \sum_i \sum_j (x^2 - \bar{x})(x_j - \bar{x}) S^2$$

pridanie systematickej neistoty pre y hodnoty nemá na smernicu (m) vplyv, pre konštantný člen (c), systematická neistota prispieva na posun dodatočným členom, ktorý je rovný práve S^2

χ^2 rozdelenie - Rozdelenie pravdepodobnosti pre χ^2 $P(\chi^2; n) = \frac{2^{-n/2}}{\Gamma(n/2)} \chi^{n-2} e^{-\chi^2/2}$

$\Gamma(x)$ je štandardná gama funkcia $\Gamma(1) = 1$ $\Gamma(1/2) = \sqrt{\pi}$ $\Gamma(n+1) = n\Gamma(n)$

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \ln^n t dt$$

pre $n = 0, 1, 2, \dots$ $\Gamma(n+1) = n!$

pre $n = 1/2, 3/2, 5/2, \dots$ $\Gamma(n+1) = n(n-1)(n-2)\dots(3/2\sqrt{\pi})$

n je počet stupňov volnosti = N - počet premenných, ktoré treba nastaviť pre minimalizáciu χ^2 .

χ^2 rozdelenie má strednú hodnotu n a rozptyl $2n$

χ^2 na stupeň volnosti by malo byť okolo jednotky - z centrálnej limitnej vety - χ^2 rozdelenie pre veľké n ide ku Gaussu (n musí byť extrémne veľké). Lepšie použiť $\sqrt{2\chi^2}$ - rozdelenie je rozumnejšou aproximáciou gaussu so strednou hodnotou $\sqrt{2n-1}$ a jednotkovým rozptylom pre hodnoty n väčšie ako ~ 30

Ak x aj y sú zaťažené chybou merania, pre jednoduchosť rovnakými s hodnotou σ a fitujeme priamkou pre $y = mx + c$ a rovnake chyby

$$\hat{m} = A \pm \sqrt{A^2 + 1} \quad \text{kde} \quad A = \frac{V(y) - V(x)}{2 \text{cov}(x, y)}$$

dve riešenia pre smernicu sú vzájomne kolmé, jedno dáva najlepšiu priamku a druhé najhoršiu. Znamienko plus je pre kladnú kovarianciu minus pre zápornú

Ak chyby nie sú rovnaké

$$\hat{m} = \frac{\sigma_y}{\sigma_x} \left(A \pm \sqrt{A^2 + 1} \right) \quad A = \frac{\sigma_x^2 V(y) - \sigma_y^2 V(x)}{2\sigma_x \sigma_y \text{cov}(x, y)}$$

Ak sú chyby rôzne pre každý bod - niet analytického riešenia - len numericky

Nelineárna metóda najmenších štvorcov

vo všeobecnosti sa musí použiť iteratívna technika. Urobí sa prvý odhad \mathbf{a}^0 , potom gradienty sú (pre nezávislé merania). V iteračnom riešení - dobrý prvý odhad reprezentuje 90% procesu. Vytvoríme matice a invertujeme. Keď g_r sa stane dostatočne malé, iteráciu zastavíme a výsledné \mathbf{a} je riešením - výhodné je využitie softwarových balíkov.

6 PRAVDEPODOBNOŠŤ A SPOĽAHLIVOSŤ (PROBABILITY AND CONFIDENCE)

UČEBNÉ CIELE

Cieľom tejto kapitoly je oboznámiť študentov s pojmom pravdepodobnosti, s rôznymi definíciami pravdepodobnosti, čo je to matematická, empirická, objektívna, subjektívna pravdepodobnosť (tiež známa ako Bayesova štatistika), čo vyjadrujú intervaly spoľahlivosti (binomický, poissonovský, gaussovský), ako je možné využiť Studentove t rozdelenie.

KLÚČOVÉ SLOVÁ

Matematická, empirická, objektívna pravdepodobnosť, Bayesova štatistika, úroveň spoľahlivosti, intervaly spoľahlivosti, Studentove t rozdelenie.

6.1 PRAVDEPODOBNOŠŤ

Môžeme sa stretnúť so 4 definíciami pravdepodobnosti - matematická, empirická, objektívna, subjektívna

6.1.1 Matematická pravdepodobnosť

Nech $S = \{E_1, E_2, \dots\}$ je súbor možných výsledkov experimentu (eventy). Eventy sú vzájomne sa vylučujúce. Pre každý event E je pravdepodobnosť $P(E)$ reálnym číslom, vyhovujúce **axiómam pravdepodobnosti**:

I. $P(E) \geq 0$

II. $P(E_1 \text{ alebo } E_2) = P(E_1) + P(E_2)$ pre vzájomne sa vylučujúce javy

III. $\sum P(E_i) = 1$ kde sa sumuje cez všetky vzájomne sa vylučujúce eventy

Tento prístup formuloval A.N. Kolmogorov, uvádzané axiómy sú zjednodušeným vyjadrením jeho prístupu. Napr. z axiómu III vyplýva - $P(\text{nie } E) = 1 - P(E)$ a teda $P(E) \leq 1$

6.1.2 Empirická pravdepodobnosť

Je to "ortodoxná" definícia - experiment je realizovaný N krát a a istý výstup A je pozorovaný v M prípadoch. Keď N ide do nekonečna, pomer M/N smeruje k limite, ktorá je definovaná ako pravdepodobnosť $P(A)$ javu A . N opakovaní môže byť realizovaných opakovaním jedného experimentu N krát, alebo súčasnou realizáciou N rovnakých experimentov.

problémy - pravdepodobnosť eventy nie je vlastnosťou čiastkového určitého eventy, ale vlastnosťou experimentu a celého súboru.

- experiment musí byť opakovateľný v rovnakých podmienkach, s rôznymi možnými výstupmi

6.1.3 3. Objektívna - sklon

Pravdepodobnosť sa chápe ako objektívna veličina s vlastnými pravidlami (prístup znovu vskriesený v druhej polovici minulého storočia - sir Karl Popper. Jeho argument bol že, podľa kvantovej mechaniky, veda predpovedá pravdepodobnosti nie istoty. Ak pravdepodobnosť nie je objektívna realita, pretože jej hodnota závisí na súbore, vybranom na základe ľubovôle počtára, potom ani častice, ktoré popisuje, nemajú "reálne" vlastnosti, správanie alebo existenciu. Popper preto navrhuje objektívnu pravdepodobnosť, ktorá existuje so svojimi vlastnými zákonmi - i keď jej jediný pozorovateľný efekt je normálna frekvenčne limitovaná pravdepodobnosť.)

objektívna pravdepodobnosť vyzerá rozumná vo vhodných prípadoch (keď sa uvažujú presne rovnaké pravdepodobnosti) - perfektná minca musí mať 50% pravdepodobnosť, že padne číslo alebo znak, symetrická kocka má rovnaké pravdepodobnosti pre každú stranu (= 1/6). Ak sa zaoberáme spojitými premennými vznikajú problémy - takýto prístup zlyháva.

6.1.4 Subjektívna pravdepodobnosť (Bayesova štatistika)

Subjektívny prístup k pravdepodobnosti je tiež známy ako Bayesova štatistika - aby sme o nej mohli diskutovať potrebujeme Bayesov teorém - pre jeho objasnenie je potrebná definícia podmienenej pravdepodobnosti - $p(a|b)$ ak b je pravda potom a - pravdepodobnosť javu a ak nastal jav b .

Napr.ak vyberáme deň v týždni náhodne, potom pravdepodobnosť, že to bude utorok je $p(\text{utorok})=1/7$. Ak vieme, že je to pracovný deň (deň pracovného týždňa), dostaneme $p(\text{utorok}|\text{pracovný deň})=1/5$.

Bayesov teorém (rev. Thomas Bayes 1763)

$p(a|b)p(b) = p(a \text{ a súčasne } b) = p(b|a)p(a)$ z toho

$$p(a|b) = \frac{p(b|a)p(a)}{p(b)}$$

často je užitočné písať $p(b)$ ako pravdepodobnosť, že nastane b bez ohľadu či je a pravdivé alebo nie (\bar{a} označuje "nie a ")

$$p(b) = p(b|a)p(a) + p(b|\bar{a})[1 - p(a)]$$

napr. Čerenkovov počítač

Máme zväzok mezónov pozostávajúci z 90% piónov a 10% kaónov ktorý dopadá do Čerenkovovho počítača. V princípe dáva počítač signál len pre pióny ale nie pre kaóny, takýmto spôsobom sa identifikuje každý jednotlivý mezón. Detektor má 95 % účinnosť ak dáva signál od pionu a tiež je 6% pravdepodobnosť, že dá náhodný signál od kaónu. Ak mezón dáva signál, môžeme použiť Bayesov teorém pre pravdepodobnosť, že signál je od piónu

$$\begin{aligned} p(\pi|\text{signal}) &= \frac{p(\text{signal}|\pi)}{p(\text{signal}|\pi)p(\pi) + p(\text{signal}|K)p(K)} p(\pi) = \\ &= \frac{0.95}{0.95 \times 0.90 + 0.06 \times 0.10} \times 0.90 = 99.3\% \end{aligned}$$

pravdepodobnosť, že signál vyvolal kaón je komplementárna

$$p(K|\text{signal}) = 0.7 \%$$

ak nemáme signál, potom

$$\begin{aligned} p(K|\text{žiadny signál}) &= \frac{p(\text{signál}|\pi)}{p(\text{signál}|\pi)p(\pi) + p(\text{signál}|K)p(K)} p(\pi) = \\ &= \frac{0.94}{0.05 \times 0.90 + 0.94 \times 0.10} \times 0.10 = 67.6\% \end{aligned}$$

Prítomnosť signálu indikuje s vysokou pravdepodobnosťou pión, neprítomnosť signálu ukazuje na kaón, ale nie veľmi isto.

Subjektívna pravdepodobnosť, tiež známa ako Bayesova štatistika, posúva Bayesov teorém ďalej tým, že umožňuje ho aplikovať na tvrdenia označované v frekvenčnom prístupe ako "nevedecké". Pravdepodobnosť teórie (napr. že bude zajtra pršať, alebo že parita sa nenarúša) je uvažovaná ako "stupeň viery" ktorý možno môže byť meraný podľa toho čo nesúhlasíca osoba je ochotná staviť. Nasledujúci experimentálny dôkaz potom modifikuje počiatočný stupeň viery, zosilní ju alebo ju zoslabí - podľa toho či výsledok súhlasí alebo nesúhlasí s predpoveďou.

Matematicky to môžeme vyjadriť

$$p(\text{teória}|\text{výsledok}) = \frac{p(\text{výsledok}|\text{teória})}{p(\text{výsledok})} p(\text{teória})$$

Toto sa správa dosť citlivo. Ak je výsledok zakázaný nejakou teóriou ($p(\text{výsledok}|\text{teória}) = 0$) potom pozorovanie výsledku vyvráti teóriu t.j. $p(\text{teória}|\text{výsledok})$ je nula. Ak výsledok je predpovedaný teóriou s veľmi malou pravdepodobnosťou, potom pozorovanie výsledku znižuje stupeň dôvery v teóriu a naopak.

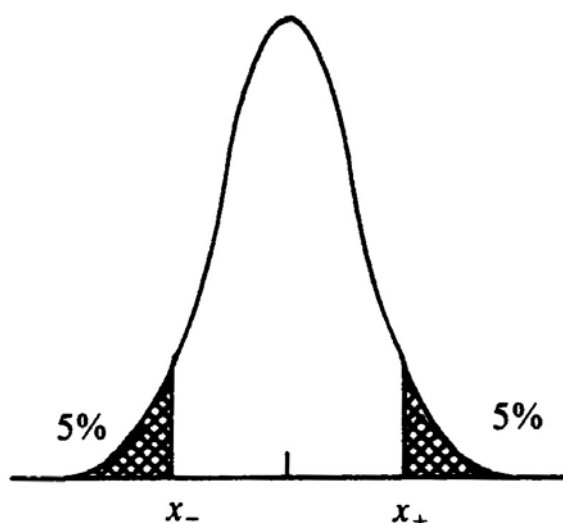
Všetko je dobré, ak počiatočný stupeň dôvery je akceptovaný. V tom je ale problém. Každý pokus založiť počiatočnú vieru (dôveru) na hádaní alebo inštinkte je nevedecký a nespoľahlivý. Jedine spôsob uvedený pre objektívnu pravdepodobnosť môže dať prijateľný počiatočný odhad stupňa dôvery, ale ten zas je nepoužiteľný pre spojitý prípad. Najčastejšie je využívaný frekvenčný prístup.

6.2 ÚROVEŇ SPOĽAHLIVOSTI (CONFIDENCE LEVEL)

Úroveň spoľahlivosti, v slovenskej literatúre tiež označovaná ako hladina významnosti, alebo konfidenčná pravdepodobnosť (čo je pojem používaný v slovenských normách), je súčasťou popisnej štatistiky, ako spôsob popisu rozšírenia rozdelenia, zvlášť v chvostoch (krajných častiach) rozdelenia.

6.2.1 Úroveň spoľahlivosti v popisnej štatistike.

Význam úrovne spoľahlivosti nám dokumentuje nasledovný príklad. Predpokladajme, že balíčky cereálií boli plnené podľa Gaussovho rozdelenia so strednou hodnotou 520 g a štandardnou odchýlkou 10g. Keď pozrieme na integrovaný Gauss (obr.6.1) - 68% balíkov bude vážiť viac než 510g a menej než 530g. Ak povieme, že váha balíkov leží v intervale 510 - 530 g, bude to správne v 68 % prípadov. Vyslovili sme tvrdenie so 68% spoľahlivosťou. Pravdepodobnosť je definovaná na základe frekvenčného prístupu ak počet balíčkov je veľký



Obr.6.1 90% centrálny interval spoľahlivosti pre Gaussove rozdelenie

Bežné hodnoty spoľahlivosti - 68% (alebo 1σ), 95.4 % (2σ), 90% (1.64σ), 95% (1.96σ) a 99 % (2.58σ) (násobok štandardnej odchýlky sa v slovenskej norme označuje ako koeficient pokrytia)

Požiadavka vysokej spoľahlivosti vedie k širokému intervalu - v praxi je bežná úroveň 90, 95%, u perfekcionista 99%.

Pre negaussovské rozdelenia - vzťah medzi úrovňou spoľahlivosti a σ neplatí.

Sú 3 konvenčné cesty výberu limitov intervalov okolo centra. Ak označíme limity x_- a x_+ , pre danú úroveň spoľahlivosti máme:

$$Pravdep(x_- \leq x \leq x_+) = \int_{x_-}^{x_+} P(x)dx = C$$

Dodatočné požiadavky sú: $x_+ - \mu = \mu - x_-$

1. symetrický interval
2. najkratší interval, limity sú také, že interval je taký krátky ako je to možné, $x_+ - x_-$ je minimálne
3. centrálnosť intervalu - pravdepodobnosť nad a pod intervalom sú rovnaké

t.j. $\int_{-\infty}^{x_-} P(x)dx = \int_{x_+}^{\infty} P(x)dx = (1 - C) / 2$

Pre symetrické rozdelenia tieto požiadavky nie sú problém

Pre jednostranné limity existuje horný a dolný limit

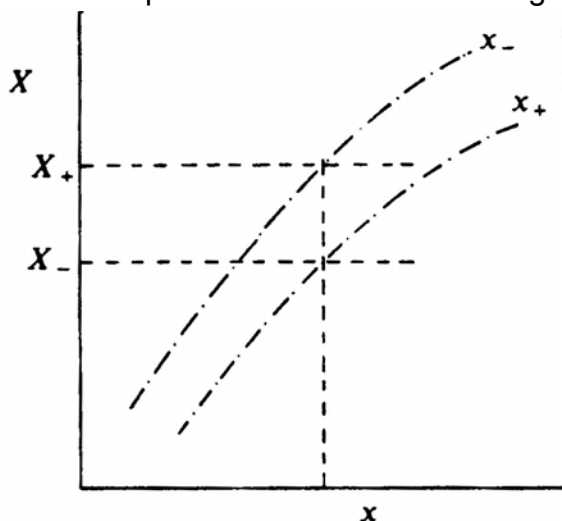
$$Pravdep(x < x_+) = \int_{-\infty}^{x_+} P(x)dx = C \quad \text{a podobne} \quad Pravdep(x > x_-) = \int_{x_-}^{\infty} P(x)dx = C$$

Treba si uvedomiť rozdiely medzi jednostranným limitom a centrálnym - 95% interval spoľahlivosti a 95% horný limit nie je to isté (prvý má 97.5% pravdepodobnosť, že obsah je pod, 2.5% že je nad, druhý 95% pod a 5% nad)

6.2.2 Intervaly spoľahlivosti v odhadoch

Predpokladajme, že chceme vedieť hodnotu parametra X a budeme ho ohodnotiť zo súboru dát dávajúcich výsledok x . Vieme presnosť nášho merania a teda $V(x)$ a štandardnú odchýlku. Problém je zahrnúť tieto znalosti do vyjadrenia intervalu spoľahlivosti.

Príklad - **nemožná pravdepodobnosť** - Prázdna nádoba váži 25.30 ± 0.14 g. Nádoba s práškom váži 25.50 ± 0.14 g. Odčítaním a kombináciou neistôt dostaneme



váhu prášku 0.20 ± 0.20 g. Naivné tvrdenie hovorí, že je 32% šanca, že váha bude viac ako 1σ od priemeru - t.j. po rozdelení je 16% šanca, že váha bude záporná - čo je absurdné tvrdenie.

Ako rozriešiť problém? Vychádzajme z frekvenčnej definície pravdepodobnosti. Pre danú hodnotu X je pravdepodobnosť získať hodnotu x daná z funkcie rozdelenia hustoty pravdepodobnosti $P(x;X)$ - pre konvenčné meranie rozlíšenia σ je to Gauss pre x so strednou hodnotou X a štandardnou odchýlkou σ . Centrálny interval

Obr. 6. 2 Diagram spoľahlivosti

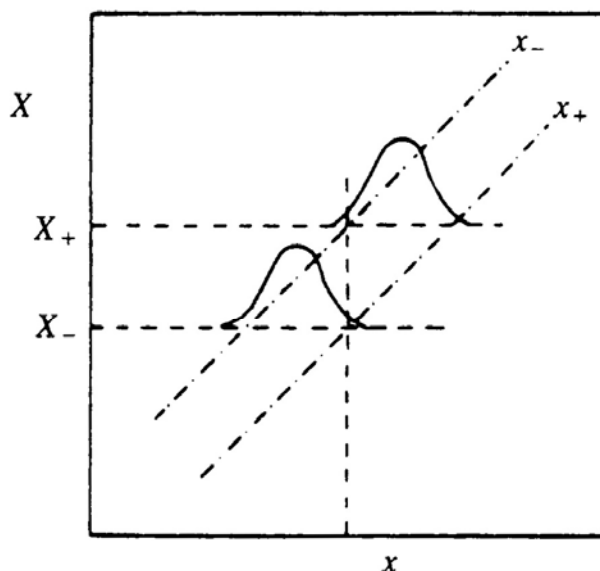
spoľahlivosti 90% - pre určitú hodnotu reálneho X hodnota merania x bude ležať (s 90% pravdepodobnosťou) v oblasti od x_- do x_+ (obr.6.1). Pre iné X to budú iné limity. Tak hodnoty x_- a x_+ môžeme uvažovať ako funkcie X (obr. 6.2). Oblasť medzi krivkami sa nazýva pás spoľahlivosti (confidence belt). Podstata týchto grafov je, že sú konštruované horizontálne - použijeme funkciu rozdelenia pravdepodobnosti $P(x;X)$ a čítame vertikálne - keď máme meranie.

Urobme aktuálne meranie x - potom x_- krivka dáva hodnotu X pre ktorú x je príslušné krivke x_- . Takto je vytvorený horný limit X_+ - neznamená to však, že X má 5% pravdepodobnosť presiahnutia X_+ - znamená to, že ak reálna hodnota X je X_+ alebo väčšia, potom pravdepodobnosť, že budeme mať meranie menšie než táto hodnota je 5% alebo menej. Podobne - hodnota X pre ktorú naša hodnota x je x_+ je dolný limit X_- . Tak sme stanovili 90% interval spoľahlivosti pre skutočnú hodnotu X ako rozpätie od X_- po X_+ . Keď sme X_- a X_+ vytvorili takouto cestou - môžeme povedať, že skutočná hodnota X leží v intervale $X_- \leq X \leq X_+$ s 90% pravdepodobnosťou. To vyzerá ako tvrdenie o X ale s skutočnosťou sa to týka X_- a X_+ .

6.2.3 Úrovne spoľahlivosti z Gaussovo rozdelenia

Pre Gaussove rozdelenie takáto konverzia z horizontálneho na vertikálne je veľmi jednoduchá. Máme meranie x s priemerom X a vieme σ - potom 90% interval spoľahlivosti pre X požaduje hodnoty X_- a X_+ také, že

$$\int_x^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x'-X_-)^2/2\sigma^2} dx' = 0.05 = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-(x'-X_+)^2/2\sigma^2} dx'$$



Obr.6.3 Diagram spoľahlivosti pre Gaussove rozdelenie

(obr.6. 3). Vzťah pre X_{\pm} požaduje, aby x ležalo nejaký počet štandardných odchýliek nad X_{\pm} . To je to isté, ako povedať, že X_{\pm} musí ležať nejaký počet σ pod meraným x čo môžeme písať v tvare

$$\int_{-\infty}^{X_{\pm}} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x'-x)^2/2\sigma^2} dx' = 0.05, \quad \text{takže interval spoľahlivosti pre Gaussovské}$$

estimátory môžeme nájsť pomocou obvyklých tabuliek Gaussových integrálov. Krivky na obr. 6.2 sa stávajú priamkami (obr.6.3) s jednotkovým gradientom, $x_{\pm} = X \pm n\sigma$ keď konštruujeme horizontálne, $X_{\pm} = x \pm n\sigma$ keď čítame vertikálne, kde n (koeficient pokrytia) je 1 pre 68% spoľahlivosť, 1.64 pre 90% spoľahlivosť, ako to uvádzajú tabuľky integrovaných gausiánov. Interval spoľahlivosti pre X získané z merania x je $x \pm n\sigma$

6.2.4 Binomický interval spoľahlivosti

Pre binomické rozdelenie pozorovaná premenná (r) je diskretná, zatiaľ čo "skutočná" hodnota (R) je spojitá. Pre diskretné premenné integrály prechádzajú na sumy - hodnota r leží v rozsahu od r_{-} po r_{+}

Keď chceme vytvoriť napr. 95% centrálny interval spoľahlivosti pre dané R , pre istotu treba r_{+} vybrať tak, že $\sum_0^{r_{+}} P(r; R) \geq 0.975$ a podobne pre r_{-} . To znamená, že naše konečné závery budú pravdivé najmenej na úrovni 95% a možno viac.

Dve krivky na diagrame spoľahlivosti budú mať schodovitý charakter pretože horizontálna súradnica je diskretná. Limity spoľahlivosti sa môžu vytvoriť z pásu ako predtým, s tým, že integrály nahradia sumy. Teda ak je nájdených m úspechov v n binomických pokusoch, limity pre individuálnu pravdepodobnosť p sú dané určením p_{-} a p_{+} nasledovne (použitím 95% centrálného intervalu ako v príklade)

$$\sum_{r=m+1}^n P(r; p_{+}, n) = 0.975 \quad \sum_{r=0}^{m-1} P(r; p_{-}, n) = 0.975$$

Tieto limity sú známe ako Clopper-Pearsonove limity spoľahlivosti.

Príklad - binomický interval spoľahlivosti.

- Vo vzorke je 20 prskaviek, 4 sú zvlhnuté. Aký je 95% limit spoľahlivosti na zastúpenie zvlhnutých prskaviek?

Riešenie

dolný limit je daný $\sum_{r=0}^3 P(r; p_{-}, 20) = 0.975$

Pokusy a chyby ukazujú, že pre $p = 0.057$, pravdepodobnosti od 0 do 3 úspechov sú 0.307, 0.373, 0.216 a 0.079 čo dáva sumu 0.975.

Horný limit je daný $\sum_{r=5}^{20} P(r; p_{+}, 20) = 0.975$ s čím je výhodnejšie pracovať ako doplnkovou

pravdepodobnosťou $\sum_{r=0}^4 P(r; p_{+}, 20) = 0.025$

Pre $p = 0.437$ pravdepodobnosti od 0 do 4 úspechov sú 0.00001, 0.0002, 0.001 a 0.005 a 0.018, čo dáva sumu 0.025. Takže hľadané limity s 95% spoľahlivosťou sú 0.057 a 0.437

6.2.5 Poissonovské intervaly spoľahlivosti

Ak je n eventov pozorovaných z Poissonovského procesu s neznámou strednou hodnotou N , 90% horný limit (napr.) je hodnota N_+ daná nasledovne

$$\sum_{r=n+1}^{\infty} P(r; N_+) = 0.90 \quad \text{alebo ekvivalentne} \quad \sum_{r=0}^n P(r; N_+) = 0.10$$

To znamená - ak skutočná hodnota N je naozaj N_+ , pravdepodobnosť získania výsledku hodnoty n ktorá je menšia ako toto je len 10% a pre N väčšie než N_+ je to ešte menšie. Tak môžeme povedať, že sme na "90% spoľahliví", že N nie je väčšie ako N_+ a priemerujúc cez mnohé také tvrdenia - budeme mať pravdu v 9 prípadoch z 10.

Rovnako pre dolný limit $\sum_{r=0}^{n-1} P(r; N_-) = 0.90$

Tieto rovnice pre N_- a N_+ ktoré sú reálne čísla, nie integer) sa môžu riešiť iteráciami. Niektoré sú v nasledujúcej tabuľke.

Pre viac premenných sa zavádza pojem oblasti spoľahlivosti v parametrickom priestore

NEKTORÉ POISSONOVSKÉ LIMITY

	Horné			Dolné		
	90%	95%	99%	90%	95%	99%
$n = 0$	2.30	3.00	4.61	—	—	—
1	3.89	4.74	6.64	0.11	0.05	0.01
2	5.32	6.30	8.41	0.53	0.36	0.15
3	6.68	7.75	10.05	1.10	0.82	0.44
4	7.99	9.15	11.60	1.74	1.37	0.82
5	9.27	10.51	13.11	2.43	1.97	1.28
6	10.53	11.84	14.57	3.15	2.61	1.79
7	11.77	13.15	16.00	3.89	3.29	2.33
8	12.99	14.43	17.40	4.66	3.98	2.91
9	14.21	15.71	18.78	5.43	4.70	3.51
10	15.41	16.96	20.14	6.22	5.43	4.13

6.2.6 Studentove t rozdelenie

Keď robíme merania so známym rozlíšením - napr. váženie guľky (13.5g) váhami so známou presnosťou (napr. 0.1g) získame výsledok 13.5 ± 0.1 g ktorý sa interpretuje ako gaussovský interval spoľahlivosti.

Ak rozlíšenie nevieme - musíme urobiť viacero meraní a pozerat' na rozmazanie hodnôt. Jedno meranie dá odhad, ale nič nehovorí o presnosti, σ nie je známe a priori, ale môže byť odhadnuté zo súboru viacerých hodnôt - nedostaneme skutočnú hodnotu σ , ale len jej odhad. Ak μ je známe, použijeme vzťah

$$\hat{\sigma} = \sqrt{(x - \mu)^2} \quad (6.1)$$

ak je neznáme

$$\hat{\sigma} = s = \sqrt{\frac{N}{N-1} (x - \mu)^2} \quad (6.2)$$

Druhý prípad je obvyklejší, ale nie je univerzálny

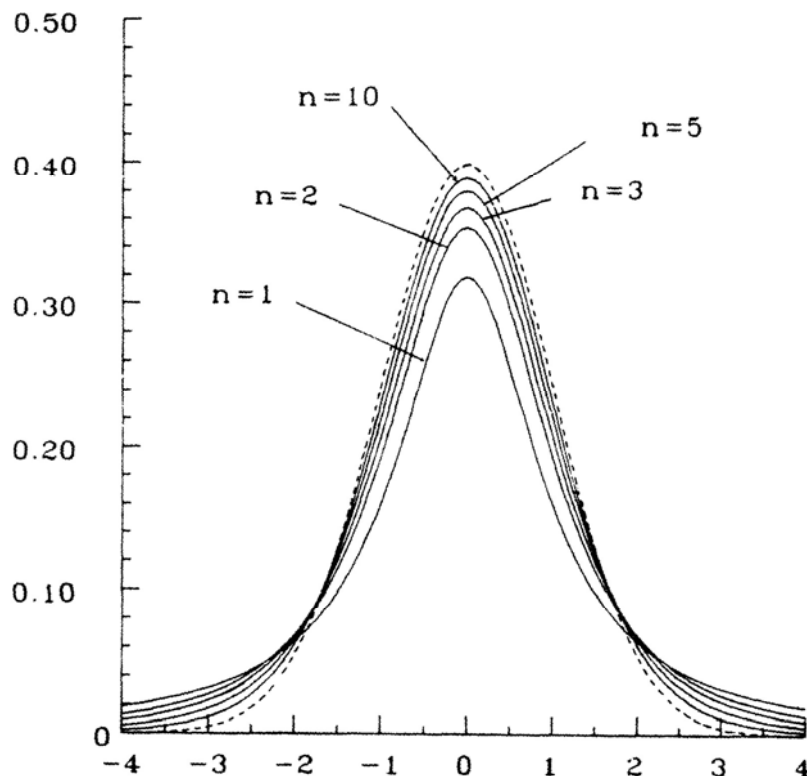
Namiesto premennej $(x - \mu)/\sigma$, ktorá je rozdelená podľa jednotkového Gaussa (t.j. má strednú hodnotu 0 a štandardnú odchýlku jednotkovú), budeme používať

$$\text{premennú } t = \frac{x - \mu}{\hat{\sigma}}$$

t nemá normálne rozdelenie s jednotkovým rozptylom, ako by to bolo ak $\sigma = \hat{\sigma}$, rozdiel je v dodatočnej neistote. V praxi, zvlášť pre malé N , odhad σ býva dosť nepresný. t je popísané Studentovým rozdelením (zavedené - William Gosset, ktorý písal pod pseudonymom "Student") t obsahuje len pozorovateľné veličiny x a $\hat{\sigma}.n$, počet stupňov volnosti v χ^2 rozdelení je N pre (rov.(6.1)) alebo $N - 1$ ak použijeme (rov 6.2). Rozdelovacia funkcia pre t je

$$f(t; n) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} \frac{1}{\left(1 + (t^2/n)\right)^{(n+1)/2}}$$

Krivka vyzerá ako Gassián, skutočne pre veľké n smeruje k jednotkovému Gaussu, ale chvosty na bokoch sú väčšie, najmä pre malé n .



Obr. 6.4 Studentove t rozdelenie pre $n= 1, 2, 3, 5$ a 10 v porovnaní s jednotkovým Gaussom (prerušovaná krivka)

Rozdiel medzi Studentovým t rozdelením a normálnym rozdelením sa stáva významným pre malé n a veľké odchýlky od centra. Použitie v praxi je komplikovanejšie ako u Gausa kvôli parametru n .

KRITICKE HODNOTY t
pre rôzne hodnoty úrovne spoľahlivosti a n

Spôľahlivosť (obojsstranná)	60%	80%	90%	95%	98%	99%
	(jednostranná) 80%	90%	95%	97.5%	99%	99.5%
$n = 1$	1.376	3.078	6.314	12.706	31.820	63.651
2	1.061	1.886	2.920	4.303	6.965	9.925
3	0.978	1.638	2.353	3.182	4.541	5.841
4	0.941	1.533	2.132	2.776	3.747	4.604
5	0.920	1.476	2.015	2.571	3.365	4.032
6	0.906	1.440	1.943	2.447	3.143	3.707
7	0.896	1.415	1.895	2.365	2.998	3.499
8	0.889	1.397	1.860	2.306	2.896	3.355
9	0.883	1.383	1.833	2.262	2.821	3.250
10	0.879	1.372	1.812	2.228	2.764	3.169
11	0.876	1.363	1.796	2.201	2.718	3.106
12	0.873	1.356	1.782	2.179	2.681	3.055
13	0.870	1.350	1.771	2.160	2.650	3.012
14	0.868	1.345	1.761	2.145	2.624	2.977
15	0.866	1.341	1.753	2.131	2.602	2.947
16	0.865	1.337	1.746	2.120	2.583	2.921
17	0.863	1.333	1.740	2.110	2.567	2.898
18	0.682	1.330	1.734	2.101	2.552	2.878
19	0.861	1.328	1.729	2.093	2.539	2.861
20	0.860	1.325	1.725	2.086	2.528	2.845
21	0.859	1.323	1.721	2.080	2.518	2.831
22	0.858	1.321	1.717	2.074	2.508	2.819
23	0.858	1.319	1.714	2.069	2.500	2.807
24	0.857	1.318	1.711	2.064	2.492	2.797
∞	0.842	1.282	1.645	1.960	2.326	2.576

Význam použitia Studentovho t smeruje k problémom súborov s nejakým priemerom \bar{x} a vzorkovým rozptylom s keď sa zaujímate o to, čo vieme o skutočnej strednej hodnote μ . To zahŕňa použitie dodatočného faktora \sqrt{n} , čo sa ukazuje vhodnejšie než len chybu σ meraní. V takom prípade $(\bar{x} - \mu) / (\sigma / \sqrt{n})$ je jednotkový Gauss, takže zodpovedajúce t bude

$$t = \frac{(\bar{x} - \mu) / (\sigma / \sqrt{n})}{\hat{\sigma} / \sigma} = \frac{\bar{x} - \mu}{\hat{\sigma} / \sqrt{n}}$$

Ohodnotíme a využijeme tabuľky.

Príklad

Štyri hodnoty {3.9, 4.5, 5.5, 6.1} sú vybraté zo suboru, ktorého stredná hodnota je známa a je 4.9. Je získaná ďalšia hodnota 7.3 - Je pravdepodobné, že patrí do súboru?

Riešenie

$\hat{\sigma}$ vypočítame ako 0.86. Odchýlka bodu pri 7.3 je $2.8\hat{\sigma}$, takže t je 2.8. Z tabuľky pre $n = 4$ pravdepodobnosť takej odchýlky je 95%. To je pravdepodobne akceptovateľné, je tu významnosť rozdielu dvoch štandardných odchýlok - s takou spoľahlivosťou musíme uvažovať.

Príklad: Intervaly spoľahlivosti pomocou Studentovho rozdelenia.

Test 25 profesorov ukazuje priemerné IQ 128 s odchýlkou $s = 15$. Aký sú 95% limity spoľahlivosti na skutočnú hodnotu priemerného IQ všetkých profesorov?

Riešenie

Predpokladajme, že týchto 25 je správne vybraná náhodná vzorka, odhadnutá chyba priemeru je $15/\sqrt{25}$. Ak je to Gauss, použijeme $\pm 1.96\sigma$ a dostaneme limity od 122.1 do 133.9. Namiesto toho použijeme Studentovo t : kritické t pre 24 stupňov voľnosti je 2.06 (miesto 1.96 u Gausa). Takže limity spoľahlivosti sú 121.8 a 134.2 - trochu širšie ako u Gausa.

ÚLOHY

6. 1 Ukážte, že ak 8 pokusov vytvorí 4 úspechy a 4 neúspechy, 90% interval spoľahlivosti pre vlastnú jednotlivú pravdepodobnosť p je od 19 do 81%. Porovnajte tieto výsledky s odhadom z Gaussovskej aproximácie.

6. 2 Pri štúdiu rozpadu protónov (zriedkavý proces) bolo pozorovaných 7 prípadov za 1 rok vo vzorke 1000000 kg vodíka. Určte 95% úroveň spoľahlivosti pre počet rozpadov a tak pre polčas premeny protónu za predpokladu, že neuvažujeme pozadie.

6. 3 Riešte problém uvedený v 2. ale za predpokladu, že očakávané pozadie z náhodných procesov je 3 eventy za rok.

6. 4 Obvyklý obsah kovu v rude je 5.6%. Vzorky rudy z možných banských oblastí ukazujú obsah kovu 6, 7, 9 a 8%. Priemerný obsah je lepší ako obvyklý, ale je to významné?

KONTROLNE OTAZKY

54. Ako môžeme definovať pravdepodobnosť?
55. V čom spočíva Bayesov prístup k pravdepodobnosti ?
56. Čo vyjadruje úroveň spoľahlivosti ?
57. Čomu zodpovedá úroveň spoľahlivosti 68% u gaussovského rozdelenia ?
58. Aký je rozdiel medzi jednostranným a obojstranným limitom pre gaussovské rozdelenie ?
59. V čom je špecifikum Studentovho t rozdelenia ?

SÚHRN

Pravdepodobnosť - môžeme sa stretnúť so 4 definíciami pravdepodobnosti - matematická, empirická, objektívna, subjektívna

Matematická pravdepodobnosť - Nech $S = \{E_1, E_2, \dots\}$ je súbor možných výsledkov experimentu (eventy). Eventy sú vzájomne sa vylučujúce. Pre každý event E je pravdepodobnosť $P(E)$ reálnym číslom, vyhovujúce **axiómam pravdepodobnosti**:

I. $P(E) \geq 0$, II. $P(E_1 \text{ alebo } E_2) = P(E_1) + P(E_2)$ pre vzájomne sa vylučujúce javy,

III. $\sum P(E_i) = 1$ kde sa sumuje cez všetky vzájomne sa vylučujúce eventy

Empirická pravdepodobnosť - "ortodoxná" definícia - experiment je realizovaný N krát a a istý výstup A je pozorovaný v M prípadoch. Keď N ide do nekonečna, pomer M/N smeruje k limite, ktorá je definovaná ako pravdepodobnosť $P(A)$ javu A . N opakovaní môže byť realizovaných opakovaním jedného experimentu N krát, alebo súčasťou realizáciou N rovnakých experimentov.

Objektívna – sklon - pravdepodobnosť sa chápe ako objektívna veličina s vlastnými pravidlami (prístup znovu vskriesený v druhej polovici minulého storočia - sir Karl Popper). Objektívna pravdepodobnosť vyzerá rozumná vo vhodných prípadoch (keď sa uvažujú presne rovnaké pravdepodobnosti). Ak sa zaoberáme spojitými premennými vznikajú problémy - takýto prístup zlyháva.

Bayesov teorém (rev. Thomas Bayes 1763)

$$p(a|b)p(b) = p(a \text{ a súčasne } b) = p(b|a)p(a) \quad \text{z toho} \quad p(a|b) = \frac{p(b|a)p(a)}{p(b)}$$

často je užitočné písať $p(b)$ ako pravdepodobnosť, že nastane b bez ohľadu či je a pravdivé alebo nie (\bar{a} označuje "nie a ")

$$p(b) = p(b|a) p(a) + p(b|\bar{a}) [1 - p(a)]$$

Úroveň spoľahlivosti - Keď pozrieme na integrovaný Gauss (obr.6.1) - 68% eventov bude v intervale stredná hodnota \pm štandardná odchýlka. Ak povieme, že event leží v intervale stredná hodnota \pm jedna štandardná odchýlka, bude to správne v 68 % prípadov. Vyslovili sme tvrdenie so 68% spoľahlivosťou. Pravdepodobnosť je definovaná na základe frekvenčného prístupu ak počet eventov je veľký

Bežné hodnoty spoľahlivosti - 68% (alebo 1σ), 95.4 % (2σ), 90% (1.64σ), 95% (1.96σ) a 99 % (2.58σ) (násobok štandardnej odchýlky sa v slovenskej norme označuje ako koeficient pokrytia). Požiadavka vysokej spoľahlivosti vedie k širokému intervalu - v praxi je bežná úroveň 90, 95%, u perfekcionista 99%.

Sú 3 konvenčné cesty výberu limitov intervalov okolo centra. Ak označíme limity x_- a x_+ , pre danú úroveň spoľahlivosti máme:

$$Pravdep(x_- \leq x \leq x_+) = \int_{x_-}^{x_+} P(x)dx = C$$

Dodatočné požiadavky sú: $x_+ - \mu = \mu - x_-$

1. symetrický interval , 2. najkratší interval, limity sú také, že interval je taký krátky ako je to možné, $x_+ - x_-$ je minimálne , 3. centrálnosť intervalu - pravdepodobnosť

nad a pod intervalom sú rovnaké, t.j. $\int_{-\infty}^{x_-} P(x)dx = \int_{x_+}^{\infty} P(x)dx = (1-C)/2$

Pre symetrické rozdelenia tieto požiadavky nie sú problém. Pre jednostranné limity existuje horný a dolný limit

$$\text{Pravdep}(x < x_+) = \int_{-\infty}^{x_+} P(x) dx = C \quad \text{a podobne} \quad \text{Pravdep}(x > x_-) = \int_{x_-}^{\infty} P(x) dx = C$$

Treba si uvedomiť rozdiely medzi jednostranným limitom a centrálnym - 95% interval spoľahlivosti a 95% horný limit nie je to isté (prvý má 97.5% pravdepodobnosť, že obsah je pod, 2.5% že je nad, druhý 95% pod a 5% nad)

Pre danú hodnotu X je pravdepodobnosť získať hodnotu x daná z funkcie rozdelenia hustoty pravdepodobnosti $P(x; X)$ - pre konvenčné meranie rozlíšenia σ je to Gauss pre x so strednou hodnotou X a štandardnou odchýlkou σ . Centrálny interval spoľahlivosti 90% - pre určitú hodnotu reálneho X hodnota merania x bude ležať (s 90% pravdepodobnosťou) v oblasti od x_- do x_+ (obr.6.1). Pre iné X to budú iné limity.

Tak hodnoty x_- a x_+ môžeme uvažovať ako funkcie X (obr. 6.2). Oblasť medzi krivkami sa nazýva pás spoľahlivosti (confidence belt).

Máme meranie x s priemerom X a vieme σ - potom 90% interval spoľahlivosti pre X požaduje hodnoty X_- a X_+ také, že

$$\int_x^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x'-X)^2/2\sigma^2} dx' = 0.05 = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-(x'-X)^2/2\sigma^2} dx'$$

(obr.6. 3). Vzťah pre X_- požaduje, aby x ležalo nejaký počet štandardných odchýliek nad X_- . To je to isté, ako povedať, že X_- musí ležať nejaký počet σ pod meraným x

čo môžeme písať v tvare $\int_{-\infty}^{x_-} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x'-x)^2/2\sigma^2} dx' = 0.05$, takže interval spoľahlivosti

pre Gaussovské estimátory môžeme nájsť pomocou obvyklých tabuliek Gaussových integrálov. Krivky na obr. 6.2 sa stávajú priamkami (obr.6.3) s jednotkovým gradientom, $x_{\pm} = X \pm n\sigma$ keď konštruujeme horizontálne, $X_{\pm} = x \pm n\sigma$ keď čítame vertikálne, kde n (koeficient pokrytia) je 1 pre 68% spoľahlivosť, 1.64 pre 90% spoľahlivosť, ako to uvádzajú tabuľky integrovaných gausiánov. Interval spoľahlivosti pre X získané z merania x je $x \pm n\sigma$

Poissonovský interval spoľahlivosti - Ak je n eventov pozorovaných z Poissonovského procesu s neznámou strednou hodnotou N , 90% horný limit (napr.) je hodnota N_+ daná nasledovne

$$\sum_{r=n+1}^{\infty} P(r; N_+) = 0.90 \quad \text{alebo ekvivalentne} \quad \sum_{r=0}^n P(r; N_+) = 0.10$$

To znamená - ak skutočná hodnota N je naozaj N_+ , pravdepodobnosť získania výsledku hodnoty n ktorá je menšia ako toto je len 10% a pre N väčšie než N_+ je to ešte menšie. Tak môžeme povedať, že sme na "90% spoľahliví", že N nie je väčšie ako N_+ a priemerujúc cez mnohé také tvrdenia - budeme mať pravdu v 9 prípadoch z

10. Rovnako to platí aj pre dolný limit $\sum_{r=0}^{n-1} P(r; N_-) = 0.90$

Studentove t rozdelenie - Ak μ je známe, použijeme vzťah $\hat{\sigma} = \sqrt{(x - \mu)^2}$ (6.1)

ak je neznáme $\hat{\sigma} = s = \sqrt{\frac{N}{N-1}(x - \mu)^2}$ (6. 2). Druhý prípad je obvyklejší, ale nie je

univerzálny. Namiesto premennej $(x - \mu)/\sigma$, ktorá je rozdelená podľa jednotkového Gausa (t.j. má strednú hodnotu 0 a štandardnú odchýlku jednotkovú), budeme

používať premennú $t = \frac{x - \mu}{\hat{\sigma}}$. t nemá normálne rozdelenie s jednotkovým rozptylom, ako by to bolo ak $\sigma = \hat{\sigma}$, rozdiel je v dodatočnej neistote. V praxi, zvlášť pre malé N , odhad σ býva dosť nepresný. t je popísané Studentovým rozdelením (zavedené - William Gosset, ktorý písal pod pseudonymom "Student") t obsahuje len pozorovateľné veličiny x a $\hat{\sigma}$, počet stupňov volnosti v χ^2 rozdelení je N pre (rov.(6.1)) alebo $N - 1$ ak použijeme (rov 6.2). Rozdelovacia funkcia pre t je

$$f(t; n) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} \frac{1}{\left(1 + (t^2/n)\right)^{(n+1)/2}}. \text{ Krivka vyzerá ako Gassián, skutočne pre veľké}$$

n smeruje k jednotkovému Gaussu, ale chvosty na bokoch sú väčšie, najmä pre malé n .

