1. Teória pravdepodobnosti

1.1 Základné pojmy, symboly a ich interpretácia

Náhodný pokus – akákoľvek činnosť, ktorej výsledok nie je predurčený podmienkami, za ktorých prebieha a ktorá je neobmedzene veľakrát (aspoň teoreticky) opakovateľná za rovnakých istých podmienok (napr. hádzanie kocky, výrobný proces...)
Hodnota náhodného pokusu sa od jedného konania k druhému mení (závisí od náhody).
Výsledkom náhodného pokusu je náhodný jav.

Náhodný jav – je akékoľvek tvrdenie o výsledku náhodného pokusu , o ktorom možno po uskutočnení pokusu rozhodnúť, či pri danej realizácii pokusu je, či nie je pravdivé.

Označenie náhodných javov: A, B, C,....

Cieľ teórie pravdepodobnosti
- zaviesť mieru početnosti výskytu náhodných javov a poskytnúť pravidlá pre manipuláciu s náhodnými javmi a s mierami početnosti ich výskytu.
Náplňou teórie pravdepodobnosti je pravdepodobnostný náhodný jav.

Elementárny náhodný jav (E_i) – jav, ktorý sa za danej situácie sa nerozkladá na menšie, čiastkové javy. Nemožno ho vyjadriť ako zjednotenie viacerých javov.

Diskrétny náhodný jav – pri realizácii náhodného pokusu na určitom intervale nadobúda končený teda spočetný počet hodnôt.

Spojitý náhodný jav – pri realizácii náhodného pokusu na určitom intervale nadobúda nekonečne veľa hodnôt.

Vzťahy a operácie medzi javmi

1. Jav A je **podmnožinou** (časťou) javu B - ak pri každom pokuse, ak nastane jav A, nastane aj jav B. $A \subset B$

2. Jav C je **zjednotením** javov A a B – je to jav, ktorý spočíva vo výskyte jedného z javov A alebo B. $C = A \cup B$

3. Jav C je **prienikom** javov A a B – je to jav , ktorý spočíva v súčasnom výskate javov A a B. $C = A \cap B$

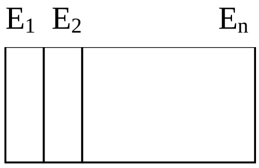
4. **Jav istý** –je jav, ktorý nevyhnutne musí nastať. Označuje sa **U**.
Jav nemožný – je jav, ktorý nemôže nastať za žiadnych okolností. Označuje sa **V**.

5. Jav $\overline{A}$ je **opačným** (protikladným, komplementárnym) javom k javu A, ak jav $\overline{A}$ nastane práve vtedy, keď nenastane jav A.
Zjednotenie opačných javov je javom istým a ich prienik je javom nemožným.

6. Javy A, B sú javy **navzájom sa vylučujúce** (**nezlúčiteľnými** alebo **disjuktnými**), ak ich prienik je javom nemožným. $A \cap B = V$

7. Javy E_1, E_2, ..., E_n tvoria **úplný súbor javov**, ak sú vzájomne disjunktné a ak ich zjednotením je jav istý

$$E_1 \cup E_2 \cup ... \cup E_n = U$$



1.2 Definície pravdepodobnosti

Pravdepodobnosť javu **-** miera opakovateľnosti výskytu výsledkov náhodného pokusu

Existujú tri prístupy definovania pravdepodobnosti náhodného javu

- klasický,
- štatistický,
- axiomatický.

1. Klasický prístup

Vychádza z toho, že množina elementárnych javov Ω obsahuje konečný počet elementárnych javov E_1, E_2, ..., E_n a že všetky tieto elementárne javy sú rovnako možné.
Tieto elementárne javy sú výsledkami určitého náhodného pokusu.

pravdepodobnosť náhodného javu A: $p(A) = \frac{m}{n}$

n – počet všetkých možných výsledkov náhodného pokusu
m – počet priaznivých výsledkov

m, n sú determinované (nemenné)

2. Štatistický prístup

Ak určitý náhodný pokus má veľké množstvo výsledkov N. Z tohto množstva môže byť náhodne vybraných n_1, n_2, n_3 ,........ prvkov, medzi ktorými sa môže vyskytovať a_1, a_2, a_3, ...priaznivých, podiel

$W(A) = \frac{a_2}{n_2}$ - je **relatívna početnosť = odhad pravdepodobnosti daného javu**

Relatívna početnosť sa tým viac približuje k pravdepodobnosti daného javu, čím viac sa rozsah výberu n_1, (n_2, n_3,....) približuje rozsahu N

3. Axiomatický prístup

Pravdepodobnosť je definovaná vetami (axiomami)

- Pravdepodobnosť náhodného javu A je nezáporné číslo, nanajvýš rovné nule : $0 \leq P(A) \leq 1$
- Pravdepodobnosť javu nemožného je rovná 0 : $P(V) = 0$
- Pravdepodobnosť javu istého je rovná 1 : $P(U) = 1$
- Ak jav A je podmnožinou javu B ($A \subset B$), potom $P(A) \leq P(B)$
- Ak jav A´ je javom opačným k javu A, ktorý má pravdepodobnosť P(A), potom pravdepodobnosť javu A´ sa vypočíta nasledovne: $P(A´) = 1 - P(A)$, čiže súčet pravdepodobností dvoch navzájom opačných javov sa rovná 1 : $P(A) + P(A´) = 1$

1.3 Analytické metódy na výpočet pravdepodobnosti zložených javov

a) Pravidlo sčítania pravdepodobnosti

(pravidlo pre výpočet pravdepodobnosti zjednotenia dvoch alebo viacerých javov)

Pravdepodobnosť zjednotenia dvoch javov sa rovná súčtu pravdepodobností týchto javov zmenšenému o pravdepodobnosť súčasného výskytu týchto javov

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (platí pre javy, ktoré sa nevylučujú)

Ak javy A, B – sú navzájom sa vylučujúce javy, potom: $P(A \cup B) = P(A) + P(B)$, pretože $A \cap B$ je javom nemožným a teda $P(A \cap B) = 0$

Pravidlo platí v uvedených tvaroch aj pre n náhodných javov.

b) Pravidlo násobenia pravdepodobnosti

(pravidlo pre výpočet pravdepodobnosti prieniku 2 alebo viacerých javov)

Pre javy závislé:
Pravdepodobnosť prieniku dvoch javov sa rovná súčinu pravdepodobnosti jedného z týchto javov a podmienenej pravdepodobnosti javu druhého.

$$P(A \cap B) = P(A) \,.\, P(B/A),$$

alebo

$$P(A \cap B) = P(B) \,.\, P(A/B),$$

Pre nezávislé javy:
Pravdepodobnosť prieniku dvoch nezávislých javov sa rovná súčinu pravdepodobností týchto javov.

$$P(A\cap B) = P(A).\, P(B)$$

Pravidlo platí aj pre n nezávislých javov.

Javy A, B sú závislé, ak: $P(B/A) \neq P(B/A')$

Javy A, B sú nezávislé, ak platí: $P(B/A) = P(B/A') = P(B)$

c) Veta o úplnej pravdepodobnosti

Predpokladajme, že jav A sa môže uskutočniť len v kombinácii s niektorým z javov B_1, B_2, ... B_n , ktoré tvoria úplný súbor javov a teda sú navzájom sa vylučujúce.
Môžu sa vyskytnúť možnosti: $A\cap B_1$ alebo $A\cap B_2$, $A\cap B_3$, ... $A\cap B_n$.
Pravdepodobnosť, že jav A sa uskutočnil, sa vypočíta ako pravdepodobnosť zloženého javu: $A = (A\cap B_1) \cup (A\cap B_2) \cup \ldots\ldots\ldots\ldots \cup P(A\cap B_n)$, čiže

$$P(A) = P(A\cap B_1)\cup P(A\cap B_2)\cup \ldots\cup P(A\cap B_n) =$$

$$= P(B_1)\,.\,P(A/B_1) + P(B_2)\,.\,P(A/B_2)\ldots\ldots\ldots = \sum_{i=1}^{n} \mathbf{P(B_i)\,.\,P(A/B_i)}$$

d) Bayesova veta (veta o pravdepodobnosti hypotézy)

Nadväzuje na vetu o úplnej pravdepodobnosti)
Predpokladáme, že jav A sa uskutočnil.
Pravdepodobnosť, že jav A sa uskutočnil práve v kombinácii s javom B_i , sa vypočíta nasledovne:

$$P(B_i / A) = \frac{P(B_i).P(A / B_i)}{\Sigma P(B_i).P(A / B_i)}$$

2. Náhodné veličiny

2.1 Charakteristika a štrukturalizácia náhodných veličín

Náhodná veličina (NV) – je to náhodná premenná, ktorá pri realizácii určitého náhodného pokusu môže nadobudnúť rôznu číselnú hodnotu v závislosti od náhody.
Náhodná veličina je výsledok náhodného pokusu vyjadrený číslom.

Hodnota náhodnej veličiny – je číselná charakteristika realizovaného náhodného pokusu a je jednoznačne určená výsledkom.

Označenie náhodných veličín : veľkými písmenami z konca abecedy napr. X, Y, Z ...

x_1, x_2, ...x_a – konkrétne hodnoty (číselné charakteristiky), ktoré náhodná veličina X nadobúda pri realizácii pokusu

Rozdelenie NV :

a) Podľa počtu položiek, ktorými sú vyjadrené:
NV jednorozmerné – sú vyjadrené len jedným číslom X,Y,Z, ... (x, y, z, ...)
NV viacrozmerné – sú vyjadrené viacerými číslami, predstavujú n-rozmerný vektor
$\overline{X}$ $x_1, x_2, \dots x_a$; $\overline{Y}$ $y_1, y_2, \dots y_a$
Z viacrozmerných sa najčastejšie sa vyskytujú dvojrozmerné NV
Dvojrozmerná NV Z = (x_i, y_j) je tvorená dvomi jednorozmernými NV

b) Podľa charakteru
NV diskrétne – môžu nadobúdať len končený (spočetný) počet hodnôt
(napr. počet nepodarkov)
NV spojité - aj na malom intervale môžu nadobúdať nekonečne veľa hodnôt
(napr. určitý parameter, rozmer,…)

2.2 Štatistické súbory a ich spracovanie

Štatistické súbory (ŠS) sa vytvárajú pri rôznych štatistických skúmaniach (empirické ŠS). Môžu sa skladať zo štatistických jednotiek najrôznejšieho druhu. Na štatistických jednotkách možno skúmať veľa rôznych štatistických znakov.

ŠS je súbor, skladajúci sa z *n* jednotiek, na ktorých sa dajú sledovať znaky X, Y, ... Jednotlivé znaky, ktoré predstavujú náhodné premenné, nadobúdajú rozličné hodnoty, napr, znak X môže nadobúdať hodnoty x_1, x_2, ... x_k s početnosťami n_1, n_2, ...n_k, kde indexy 1,2, ...k označujú jednotlivé hodnoty alebo intervaly hodnôt.

napr. 1, 2, 4, 0 ; 12,40, 12,41, 12,39, …

Usporiadaný rad hodnôt - údaje (číselné hodnoty) zoradené od najmenšej po najväčšiu:

$x_1 = x_{min}, x_2, x_3, \ldots\ldots\ x_n = x_{max}$

Rad rozdelenia početnosti

jednotlivé hodnoty	x_1	x_2	......	x_k
absolútne početnosti	n_1	n_2	......	n_k
Relatívne početnosti	p_1	p_2	......	p_k
Kumulované početnosti	n_1	n_1+n_2	.....	n

$$p_i = \frac{n_i}{n}$$

p_i - relatívna početnosť i-tej triedy, n_i - absolútna početmosť i-tej triedy,

n - celkový počet údajov v štatistickom súbore $n = \sum_{i=1}^{k} n_i$

Absolútna početnosť = počet štatistických jednotiek patriacich do určitých skupín - tried.

Relatívna početnosť = podiel štatistických jednotiek i-tej triedy na celkovom rozsahu štatist. súboru.

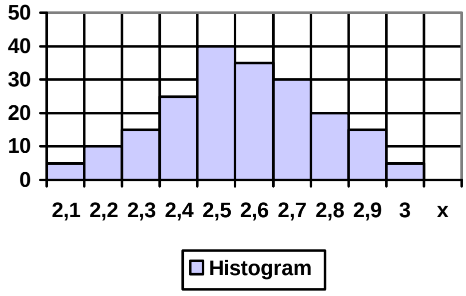
Kumulovaná početnosť = počet (resp. podiel) jednotiek ŠS, ktorý má hodnotu znaku menšiu nanajvýš rovnú hodnote danej triedy.
Túto početnosť dostaneme, ak relatívne alebo absolútne početnosti pripočítavame postupne od prvej triedy až po poslednú: n_1 $n_1 + n_2$ $n_1 + n_2 + n_3$ ...... $n_1 + n_2 + n_3 + \ldots + n_k$

Grafické znázornenie ŠS

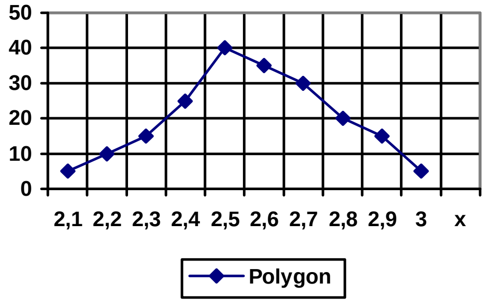
1. Histogram
2. Polygón
3. Kumulovaná krivka

Histogram rozdelenia početnosti je stĺpikový diagram, v ktorom sa na os x nanášajú jednotlivé intervaly tried a na os y absolútne alebo relatívne početnosti.

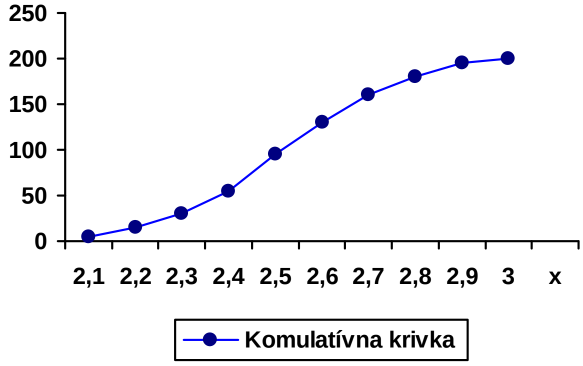
Interval	n_i
2,0 – 2,1	5
2,1 – 2,2	10
2,2 – 2,3	15
2,3 – 2,4	25
2,4 – 2,5	40
2,5 – 2,6	35
2,6 – 2,7	30
2,7 – 2,8	20
2,8 – 2,9	15
2,9 – 3,0	5
∑	200



Polygón je spojnicový diagram, ktorý vznikne pospájaním priesečníkov stredov intervalov na osi x a relatívnej (absolútnej) početnosti na osi y.



Kumulovaná (kumulatívna) krivka vznikne pospájaním priesečníkov stredov intervalov na osi x a kumulovanej početnosti na osi y.



Intervalové rozdelenie početnosti

- vznikne vytvorením systému intervalov a rozdelením údajov do príslušných intervalov

Šírka intervalov i

$$i = \frac{X_{\max} - X_{\min}}{1 + 3{,}322 \ \log n}$$

Variačné rozpätie: $R = X_{\max} - X_{\min}$

2.3 Spôsoby popísania pravdepodobnostného správania sa náhodných veličín

Pravdepodobné správanie sa náhodných veličín možno popísať dvomi spôsobmi:

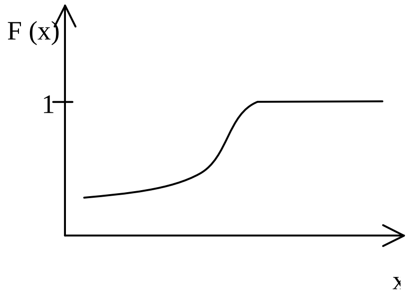
a) **Distribučná funkcia**
   - **možno použiť pri diskrétnych aj pri spojitých NV**

b) **Zákon rozdelenia pravdepodobnosti – pre diskrétne premenné (DNV)**
   Zákon rozdelenia hustoty pravdepodobnosti - pre spojité premenné (SNV)

Distribučná funkcia – je funkcia, ktorá každému x $\in$ (- ∞ , ∞) priraďuje pravdepodobnosť toho, že NV nadobudne hodnotu menšiu alebo nanajvýš rovnú x.

Pre všetky x $\in$ (- ∞, ∞)

$$F(x) = P(X \le x) = P(-\infty < X \le x)$$



Vlastnosti distribučnej DF

1. $P(x_1 \le X \le x_2) = F(x_2) - F(x_1)$
2. $F(-\infty) = 0$
3. $F(\infty) = 1$
4. $0 \le F(x) \le 1$
5. F(x) neklesajúca funkcia
6. F(x) z ľava spojitá funkcia

Zákon rozdelenia pravdepodobnosti (pre DNV)

- poznáme vtedy, keď poznáme všetky hodnoty, ktoré môže nadobudnúť NV a zároveň poznáme pravdepodobnosti týchto hodnôt.

Môže byť daný:
- tabuľkou,
- grafom,
- funkčným predpisom

Distribučná funkcia DNV sa počíta ako kumulovaná relatívna početnosť (pravdepodobnosť)

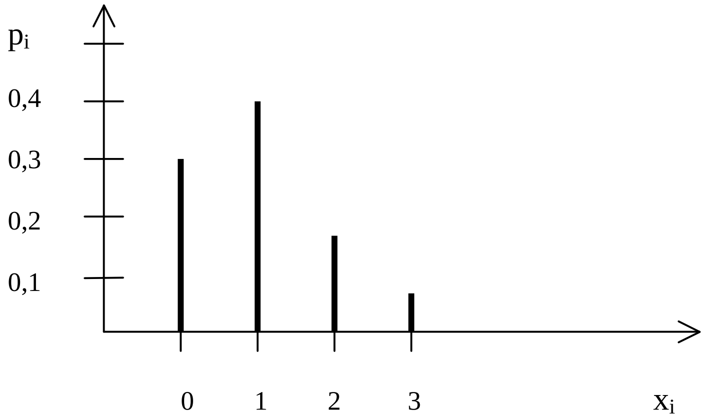
$$F(x_i) = p(x_1) + p(x_2) + ...p(x_i) = \sum_{i=1} p(x_i)$$

$$0 \le F(x) \le 1$$

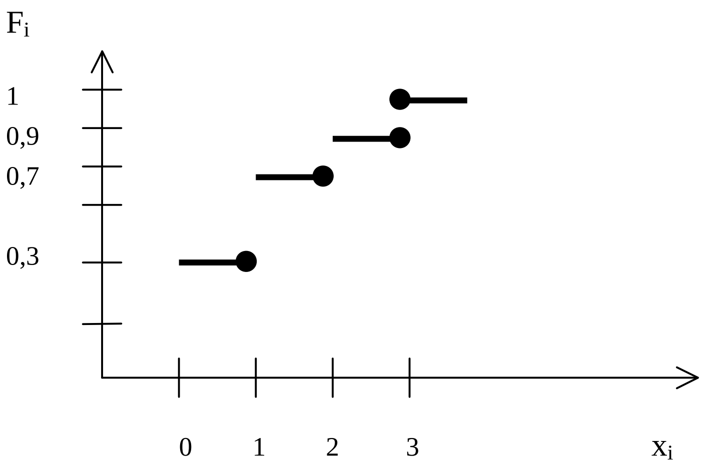
Tabuľka

x_i	0	1	2	3
p_i	0,3	0,4	0,2	0,1
$F_{(xi)}$	0,3	0,7	0,9	1

Zákon rozdelenia pravdepod. daný grafom – grafické znázornenie pravdepodobnosti



Graf distribučnej funkcie



Funkčný predpis – pravdepodobnostná funkcia

Napr. $$P(x) = \binom{n}{x} p^x q^{n-x}$$

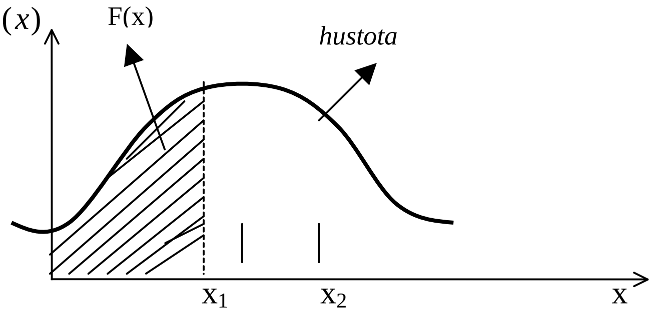
Zákon rozdelenia hustoty pravdepodobnosti (pre SNV)

Hustota pravdepodobnosti je nezáporná funkcia, ktorá pre všetky x∈(-∝,∝) vyhovuje vzťahu

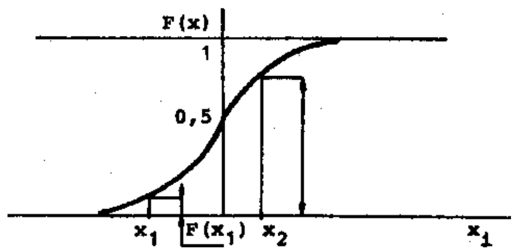
$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Hustota pravdepodobnosti: $f(x) = F'(x)$

graf hustoty pravdepodobnosti



$$P(x_1 \leq x \leq x_2) = F(x_2) - F(x_1) = \int_{-\infty}^{x_2} f(t)dt - \int_{-\infty}^{x_1} f(t)dt = \int_{x_1}^{x_2} f(t)dt$$



Dvojrozmerné náhodné veličiny

Pravdepodobnostné správanie sa dvojrozmerných resp. n-rozmerných náhodných veličín sa popisu tiež pomocou distribučnej funkcie a zákona rozdelenia pravdepodobnosti a zákona rozdelenia hustoty pravdepodobnosti.

3. Kvantitatívne charakteristiky NV

Sú to číselné hodnoty, ktoré popisujú pravdepodobnostné správanie sa NV

Možno ich rozdeliť do štyroch základných skupín:

a) charakteristiky polohy,
b) charakteristiky variability,
c) charakteristiky šikmosti,
d) charakteristiky špicatosti.

3.1 Charakteristiky polohy

Sú číselné hodnoty, ktoré charakterizujú úroveň hodnôt v štatistickom súbore

Stredná hodnota (očakávaná hodnota, matematická nádej) – označuje sa E(x)
Má charakter pravdepodobnostného priemeru

Pre DNV: $E(x) = \sum x_i . p_i = \overline{x}$

Pre SNV: $E(x) = \int_{-\infty}^{\infty} x f(x) dx$

Aritmetický priemer: $E(x) = \dfrac{\sum x_i n_i}{\sum n_i}$

Ďalšie charakteristiky polohy:

- medián $\tilde{x}$
- modus $\hat{x}$
- kvantily:

Pr.

x_i	1	2	3
p_i	0,3	0,4	0,5

Stredná hodnota uvedeného štatistického súboru :
$E_{(x)} = 1 \; 0,3 + 2 \; 0,4 + 3 \; 0,3 = 2$

Medián - prostredná hodnota usporiadaného radu hodnôt. Označuje sa Me alebo $\tilde{x}$,
Modus - hodnota s najväčšou pravdepodobnosťou výskytu (najpočetnejšia hodnota), označuje sa Mo alebo $\hat{x}$

Kvantily (Q_i) hodnoty v štatist. súbore, ktoré rozdeľujú štat. súbor v určitom pomere

Q_i - i % kvantil

Napr. Q
P ($x \leq Q_{25}$) $\geq$ 0,25 ; P ($x \geq Q_{25}$) $\leq$ 0,75

Qvartily – rozdeľujú štatistický súbor na 4 časti
Q_{25} - dolný kvartil
Q_{75} – horný kvartil
Q_{50} - prostredný kvartil = medián

Decily - súbor delia na 10 častí

Percentily – delia súbor na desať častí

3.2 Charakteristiky variability

Vyjadrujú menlivosť (premenlivosť, variabilitu) hodnôt

Disperzia D(x) = rozptyl σ^2
= stredná hodnota štvorcov odchýlok jednotlivých hodnôt x_i od strednej hodnoty E(x)

Pre DNV: $D(x) = E[x - E(x)]^2$

Pre SNV: $D(x) = \int_{-\infty}^{\infty} [x - E(x)]^2 f(x)dx$

rozptyl:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 n_i}{\sum n_i} \qquad \text{alebo} \qquad \sigma^2 = \sum (x_i - \bar{x})^2 p_i$$

D(x) = σ^2

Vlastnosti disperzie D(x):
1. D(k) = 0 k - konštanta
2. D(kX) = k^2D(X)
3. D(X ± Y) = D(X) + D(Y) Z = (X ± Y)
4. D(k ± X) = D(X)

5. Disperziu možno vypočítať pomocou momentových charakteristík

$$D(X) = \underbrace{E(x^2)} - \{\underbrace{E(x)}\}^2 = \sum x_i^2 . p_i - \left[\sum x_i p_i\right]^2 = \upsilon_2 - \upsilon_1^2$$

Smerodajná odchýlka σ :

- je odmocnina z rozptylu

$$\sigma = \sqrt{\frac{1}{n}\sum\left(x_i - \bar{x}\right)^2 n_i} \qquad \text{alebo} \qquad \sigma = \sqrt{\sum\left(x_i - \bar{x}\right)^2 p_i}$$

Variačný koeficient

- je bezrozmernou mierou variácie – udáva sa v percentách

$$V = \frac{\sqrt{D(x)}}{E(x)}.100[\%]$$

Empirický vzťah : $v = \frac{\sigma}{\bar{x}}.100[\%]$

Variačné rozpätie

- interval, v ktorom sa hodnoty pohybujú

$$R = X_{max} - X_{min}$$

Normovaná odchýlka

- vznikne transformáciou náhodnej veličiny X, ktorá má strednú hodnotu E(X) a disperziu D(X)

$$Y = \frac{X - E(X)}{\sqrt{D(X)}}$$

Stredná hodnota premennej (náhodnej veličiny) Y

E(Y) = 0

D(Y) = 1

Empirický vzťah (používaný pri riešení úloh): $t = \frac{x - \bar{x}}{\sigma}$

Momentové charakteristiky - doplnkové charakteristiky

moment - číselná charakteristika pravdepodobnostného správania sa náhodnej veličiny

- lepšie vyjadruje vplyv extrémnych hodnôt, ktorých pravdepodobnosť výskytu je veľmi malá

Pr.

X_i	1	2	3	4
p_i	0,4	0,35	0,24	0,01

$\sum P_i = 1$

počiatočné momenty:

$\nu_1 = E(x) = \sum x_i p_i = 1 . 0,4 + 2 . 0,35 + 4 . 0,24 + 10 . 0,01 = 2,16$

$\nu_2 = E(x^2) = \sum x^2_i p_i = 1^2 . 0,4 + ... = 6,64$

Skupiny momentov:

1. **všeobecné momenty** **- hodnoty okolo ľubovolnej hodnoty c**

$$\mu'_k = E\left\{(x-c)^k\right\}$$

k **-** je rád momentu

2. **počiatočné momenty,** **c = 0**

pre DNV: $\upsilon_1 = E(x) = \sum x_i p_i$ $\qquad \upsilon_1 = \dfrac{\sum x_i n_i}{n}$

$$\upsilon_2 = E(x^2) = \sum x_i^2 p_i$$
$$\upsilon_3 = E(x^3) = \sum x_i^3 p_i$$
$$\upsilon_4 = E(x^4) = \sum x_i^4 p_i$$

SNV $\quad \upsilon_k = \displaystyle\int_{-\infty}^{\infty} x^k f(x)dx$

3. **centrálne momenty** - momenty okolo strednej hodnoty, **c = E(x)**

$$\mu_k = E\left\{[X - E(X)]^k\right\}$$

$\mu_1 = \sum\left\{[X - E(X)]\right\} p_i = 0$ $\quad$ **μ_1=0 vždy!** **(1. centrálny moment = 0)**

$$D(X) = \mu_2 = E\left\{X - E(X)\right\}^2 = \sum(x_i - \bar{x})^2 p_i = \sum x_i p_i - \left(\sum x_i p_i\right)^2 = \upsilon_2 - \upsilon_1$$

$\Rightarrow$ **μ_2 = D(X)**

$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3$

$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4$

3.3 Charakteristiky šikmosti

Určujú, či rozdelenie je symetrické alebo nesymetrické (asimetrické)

Koeficient šikmosti

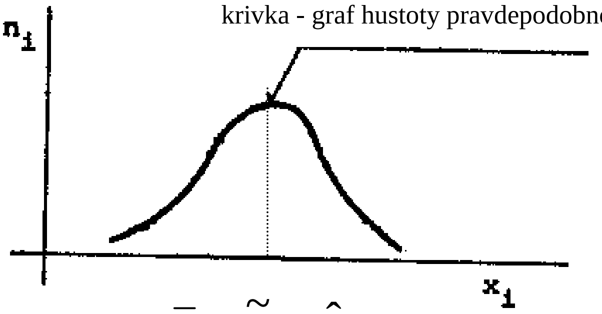
$$A = \frac{\mu_3}{\sigma^3} = \ldots\ldots \Rightarrow \ldots \begin{matrix} \langle 0 \\ 0 \\ \rangle 0 \end{matrix}$$

μ_3 - centrálny moment 3. rádu

σ^3 - smerodajná odchýlka

rozdelenie symetrické A = 0

krivka - graf hustoty pravdepodobnosti

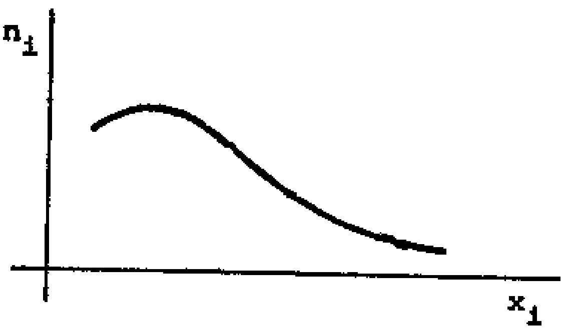


$$\bar{x} = \tilde{x} = \hat{x}$$

asymetrické rozdelenia:

A > 0

A < 0

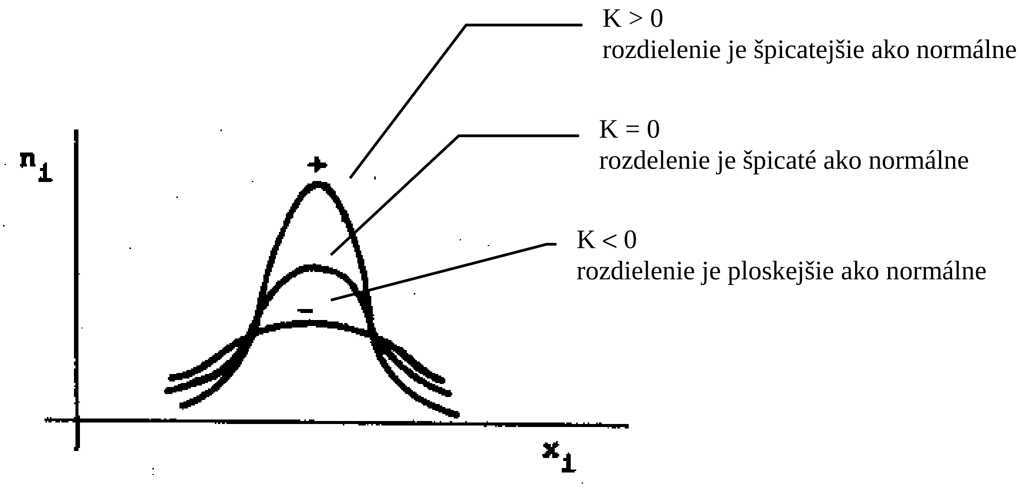


zošikmenie doľava

zošikmenie doprava

3.4 Charakteristiky špicatosti

Čím viac sú údaje sústredené okolo strednej hodnoty, tým je graf rozdelenia špicatejší.



Koeficient špicatosti: $K = \dfrac{\mu_4}{\sigma^4} - 3 \ldots\ldots\ldots = \ldots\ldots \Rightarrow \ldots\ldots \begin{cases} < 0 & \text{rozdelenie ploskejšie} \\ = 0 & \text{normálne rozdelenie} \\ > 0 & \text{špicatejšie rozdelenie} \end{cases}$

Príklad:
Vypočítajte základné kvantitatívne charakteristiky!

x_i	2	4	6	8
p_i	0,4	0,3	0,2	0,1

1. stredná hodnota: $E(x) = \nu_1$
2. smerodajná odchýlka: $\sigma = \sqrt{D(x)}$ $\quad D(x) = \mu_2 = \nu_2 - \nu_1^2$
3. koeficient šikmosti: $A = \dfrac{\mu_3}{\sigma^3}$
4. koeficient špicatosti: $K = \dfrac{\mu_4}{\sigma^4} - 3$

$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3$

$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4$

počiatočný moment: 1. rádu $\quad \upsilon_1 = E(x) = \sum x_i p_i$ **= 4 =** $\bar{x}$

$$\upsilon_1 = \frac{\sum x_i n_i}{n}$$

2. rádu $\quad \upsilon_2 = E(x^2) = \sum x_i^2 p_i$ **= 20**
3. rádu $\quad \upsilon_3 = E(x^3) = \sum x_i^3 p_i$ **= 116,8**
4. rádu $\quad \upsilon_4 = E(x^4) = \sum x_i^4 p_i$ **= 752**

1. $\upsilon_1 = E(x) = \sum x_i p_i$ **= 4 =** $\bar{x}$

2. $\sigma = \sqrt{D(x)}\, D(x) = \mu_2 = \nu_2 - \nu_1^2 = 20 - 16 = 4$

   $D(x) = 4$, $\sigma = 2$ - smerodajná odchýlka

3. $\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3 = 116{,}8 - 3.4.20 + 2.4^3 = 4{,}8$

   $$A = \frac{\mu_3}{\sigma^3} = \frac{4{,}8}{2^3} = 0{,}6$$ - rozdelenie asymetrické (zošikmené doľava)

4. $\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4 = 752 - 4.116{,}8.4 + 6.20.4^2 - 3.4^2 = 35{,}2$

   $$K = \frac{\mu_4}{\sigma^4} - 3 = \frac{35{,}2}{2^4} - 3 = -0{,}8$$ - rozdelenie ploskejšie ako normálne

Metóda vhodne zvoleného počiatku

Používa sa pre zjednodušenie výpočtu pri intervalovom rozdelení početnosti

Interval	n_i	p_i	x_i	x_i'	$\sum x_i' n_i$	$\sum x_i'^2 n_i$	$\sum x_i'^3 n_i$	$\sum x_i'^4 n_i$
14,20-14,30	10		14,25	-3				
14,30,14,40	20		14,35	-2				
14,40-14,50	30		14,45	-1				
14,50-14,60	40		**14,55**	0				
14,60-14,70	20		14,65	1				
14,70-14,80	10		14,75	2				

Výpočet základných kvantitatívnych charakteristík:

- **metóda vhodne zvoleného počiatku** $$x'_i = \frac{x_i x_0}{h}$$

  x_0 - je vhodne zvolený počiatok (14,5)
  h - šírka intervalov (0,1)
  x_i - stredy intervalov

$$x'_i = \frac{x_i x_0}{h} = \frac{14{,}25 - 14{,}55}{0{,}1} = -3$$

transformované

$$\upsilon_1 = E(x) = \sum x_i p_i \qquad \upsilon_1' = \frac{\sum x_i' n_i}{n}$$

$$\upsilon_2 = E(x^2) = \sum x_i^2 p_i \qquad \upsilon_2' = \frac{\sum x_i'^2 n_i}{n}$$

$$\upsilon_3 = E(x^3) = \sum x_i^3 p_i \qquad \upsilon_3' = \frac{\sum x_i'^3 n_i}{n}$$

$$\upsilon_4 = E(x^4) = \sum x_i^4 p_i \qquad \upsilon_4' = \frac{\sum x_i'^4 n_i}{n}$$

$$\mu'_3 = \nu'_3 - 3\nu'_1\nu'_2 + 2\nu'^3_1$$
$$\mu'_4 = \nu'_4 - 4\nu'_3\nu'_1 + 6\nu'_2\nu'^2_1 - 3\nu'^4_1$$

Disperzia: $D'(x) = \nu'_2 - \nu'^2_1$ $\qquad \upsilon_1' = \bar{x}' \rightarrow \bar{x} = x'.h + h_0$

Smerodajná odchýlka: $\sigma' = \sqrt{D'(x)}$ $\qquad \rightarrow \sigma = \sigma'.h$

Koeficient šikmosti: $A = \dfrac{\mu'_3}{\sigma'^3} = \dfrac{\mu_3}{\sigma^3}$

Koeficient špicatosti: $K = \dfrac{\mu'_4}{\sigma'^4} - 3$

$\boxed{\bar{x} = \bar{x}' h + x_0}$ - spätná transformácia

$\sigma = \sigma'.h$
$D(x) = D'(x)h^2$
$\mu_2 = D'(x) \quad \Rightarrow \quad \mu_2 = \mu_2'.h^2$
$\mu_3 = \mu_3'.h^3$
$\mu_4 = \mu_4'.h^4$

A, K - môžeme určiť z transformovaných hodnôt (nemusíme robiť spätnú transformáciu)

4. Teoretické rozdelenia náhodných veličín

Modely teoretických rozdelení pre diskrétne NV:
1. Binomické rozdelenie
2. Hypergeometrické rozdelenie
3. Poissonovo rozdelenie

Pre spojité NV
1. Normálne rozdelenie (Gauss - laplaceovo)
2. Logaritmicko - normálne
3. χ^2 (chí kvadrát) rozdelenie
4. Studentovo (T)
5. Snedecorovo (F)

Pre všetky tieto rozdelenia je definovaná pravdepodobnostná funkcia (pre DNV) resp. hustota pravdepodobnosti (pre SNV) a distribučná funkcia.
Ich hodnoty sú tabelované (uvedené v štatistických tabuľkách).

4.1 Binomické rozdelenie

Je najčastejšie sa vyskytujúce rozdelenie DNV
Nech X je jednorozmerná diskrétna NV nadobúdajúca hodnoty x_1, x_2, x_3, ...x_n
Pravdepodobnosť, že sledovaný jav nastane nech je konštantná pre všetky hodnoty
p = konšt. A pravdepodobnosť javu opačného q = 1-p
Rozdelenie pravdepodobnosti tejto NV nazývame binomickým, ak jej pravdepodobnostná funkcia (zákon rozdelenia pravdepodobnosti) je daná vzťahom:

$$P_x = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$ **pravdepodobnostná funkcia**

p, n sú parametre binomického rozdelenia

n - všetky možné výsledky pri náhodnom pokuse
p - pravdepodobnosť výskytu daného javu, **p = konšt.**
q – pravdepodobnosť javu opačného q = 1 – p

Distribučná funkcia

$$F(x) = \sum \binom{n}{x_i} p^{xi} q^{n-xi}$$

Príklad:

n = 100 $\qquad$ q = 1 - p $\qquad$ q = 1 - 0,1 = 0, 9

p = 0,1

x = 7 $\qquad P_7 = \binom{100}{7} .0{,}1^7 .0{,}9^{93}$ = 0,088

$F(7) = P(x \leq 7) = P_0 + P_1 + P_2 + \ldots. P_7$ $\qquad$ **- distribučná funkcia v bode 7**

Stredná hodnota NV, ktorá má binomické rozdelenie: $\qquad E(X) = n.p$

Disperzia: $\qquad D(X) = n.p.q$

Koeficient šikmosti: $\qquad A = \dfrac{q-p}{\sqrt{n.p.q}}$

Koeficient špicatosti $\qquad K = \dfrac{1-6.p.q}{n.p.q}$

V prevažnej väčšine prípadov je to asymetrické rozdelenie, zošikmené doľava.
Iba ak p = q, rozdelenie je symetrické.

Pre výpočet pravdepodobnosti, že premenná x, ktorá má binomické rozdelenie nadobáda hodnoty z intervalu (a,b) možno použiť normované normálne rozdelenie:

$x \in \langle a,b \rangle$

$P\left(a \langle x \langle b\right) = F(b) - F(a) = F(t_2) - F(t_1)$

normovaná odchýlka: $\qquad$ z tabuliek N(O, 1) - normálové normálne rozdelenie

$$Y = \frac{X - E(X)}{\sqrt{D(X)}}$$

$$t_1 = \frac{a - n.p}{\sqrt{n.p.q}}$$

$$t_2 = \frac{b - n.p}{\sqrt{n.p.q}}$$

4.2 Hypergeometrické rozdelenie

$\boxed{p \neq \text{konšt.}}$

Používa sa pri riešení úloh, kde sa uskutočňujú výbery bez opakovania
Parametre tohto rozdelenia sú : **S , K , s**

S – počet prvkov v základnom súbore (súbor, z ktorého sa výber uskutočňuje)
K – počet prvkov, ktoré majú určitú sledovanú vlastnosť
s - počet prvkov, ktoré sú vybrané zo základného súboru (počet prvkov vo výberovom súbore)

x – počet prvkov vo výberovom súbore, ktoré majú sledovanú vlastnosťou

Pravdepodobnosť, že vo výberovom súbore bude x prvkov s danom vlastnosťou možno vyjadriť nasledovne:

$$P(x) = \frac{\binom{K}{x}\binom{S-K}{s-x_i}}{\binom{S}{s}}$$ - pravdepodobnostná funkcia hypergeometrického rozdelenia

- Distribučná funkcia hypergeometrického rozdelenia:

$$F(x) = \sum_{x_i \le x} \frac{\binom{K}{x}\binom{S-K}{s-x_i}}{\binom{S}{s}}$$

Ak $\frac{s}{S} \le 0{,}05$ $\Rightarrow$ ak bude splnený tento predpoklad, môžeme hypergeometrické rozdelenie nahradiť binomickým, pričom platí: $p = \frac{K}{S}$, n = s

4.3 Poissonovo rozdelenie

Používa sa, ak počet výsledkov náhodného pokusu je veľmi veľký: n $\to$ ∞ a pravdepodobnosť výskytu daného javu je veľmi malá p $\to$ 0.

Je to rozdelenie tzv. zriedkavých javov.

Ide o jedno parametrické rozdelenie s parametrom λ (lamda)

$\lambda = n \, . \, p$

$$p = \frac{\lambda}{n}$$

$$q = 1 - \frac{\lambda}{n}$$

Pravdepodobnostná funkcia P(x):

$$P(x) = \frac{\lambda'^{x} e^{-\lambda}}{x!}$$

Distribučná funkcia F(x):

$$F(x) = \sum_{x_i \le x} \frac{\lambda^{x_i} e^{-\lambda}}{x!}$$

e = 2,718

Charakteristiky Poissonovho rozdelenia:

E(X) = λ

D(X) = λ

Ak sú splnené predpoklady, možno Poissonovým rozdelením nahradiť binomické rozdelenie

Predpoklady: $\lambda < 4$
$p < 0{,}1$

Príklad

Pravdepodobnosť, že výrobok je nepodarok je konštantná p = 0,1. Vypočítajte, aká bude pravdepodobnosť, že medzi 100 výrobkami bude viac ako 1 nepodarok.

p = 100, n = 100

$P_{x>1}$ - vypočítame ako pravdep. javu opačného tzn. pravdepodobnosť, že medzi 100 výrobkami bude najviac 1 nepodarok

jav opačný: $P'_{x} \le 1 = P_0 + P_1$

P_0 - vypočítame pomocou binomického rozdelenia

$$P_0 = \binom{100}{0} .0{,}01^{0} . 0{,}99^{100} = 0{,}369$$

$$P_1 = \binom{100}{1} .0{,}01^{1} . 0{,}99^{99} = 0{,}369$$

$$\boxed{P_x = \binom{n}{x} p^{x} q^{n-x}}$$

I. spôsob výpočtu

$P'_{x<1} = P_0 + P_1 = 0{,}735$

$P_{x>1} = 1 - P' = 1 - 0{,}735 = 0{,}264$

Pre výpočet možno použiť aj poissonovo rozdelenie:

$\lambda = n.p = 100 \, . \, 0{,}01 = 1$ predpoklady sú splnené

$$P_0 = \frac{1^{0} e^{-1}}{0!} = \frac{1e^{-1}}{1} = e^{-1} = \frac{1}{e} = 0{,}3678$$

$$P_1 = \frac{1^{1} e^{-1}}{1!} = \frac{1e^{-1}}{1} = \frac{1}{e}$$

$$P'_{x \le 1} = P_0 + P_1 = \frac{1}{e} + \frac{1}{e} = \frac{2}{e} = 0{,}7352$$

$$P_{x \rangle 1} = 1 - P' = 1 - 0{,}7358 = 0{,}264$$

II. spôsob výpočtu = jednoduchší a rýchlejší

4.4 Normálne rozdelenie (Gauss-Laplaceovo)

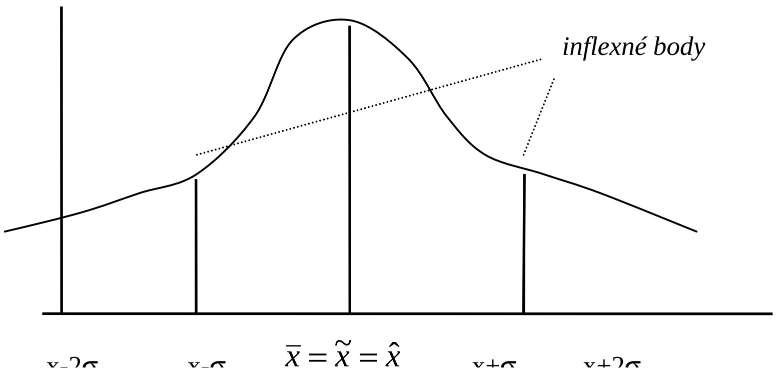
Je najčastejšie sa vyskytujúcim rozdelením spojitých náhodných veličín.

Rozdelenie pravdepodobnosti spojitej NV nazývame normálnym, ak jeho hustotu pravdepodobnosti možno zapísať vzťahom:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Parametre normálneho rozdelenia sú: $\bar{\mathbf{x}}$ **- stredná hodnota, σ - smerodajná odchýlka**

Graf hustoty pravdepodobnosti je symetrická Gaussova krivka



Distribučná funkcia:
(pri spojitej NV distribučná funkcia = integrál hustoty pravdepodobnosti):

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-\bar{x})^2}{2\sigma^2}} du$$

Pre normálne rozdelenie platí:

$$P(\bar{x} - \sigma \langle X \langle \bar{x} + \sigma) = 0{,}65$$
$$P(\bar{x} - 2\sigma \langle X \langle \bar{x} + 2\sigma) = 0{,}95$$
$$P(\bar{x} - 3\sigma \langle X \langle \bar{x} + 3\sigma) = 0{,}9973$$

Normované normálne rozdelenie

Akékoľvek normálne rozdelenie s parametrami **(x, σ)** možno pretransformovať na normované normálne rozdelenie pomocou normovanej odchýlky:

$$t = \frac{x - \overline{x}}{\sigma}$$

Normované normálne rozdelenie má strednú hodnotu rovnú 0 a smerodajnú odchýlku rovnú 1, preto sa označuje ako N(0,1)

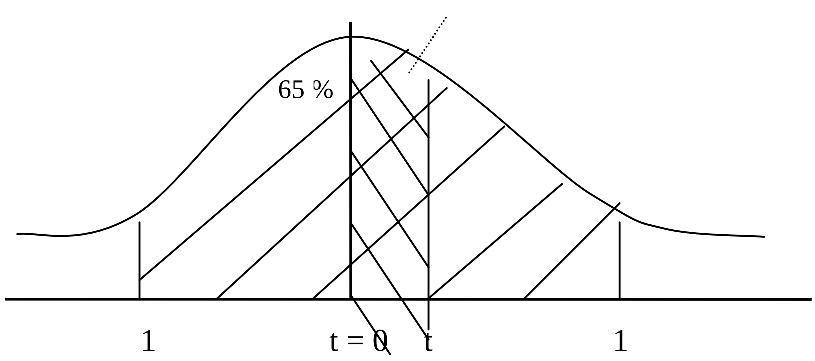
Hustota pravdepodobnosti normovaného normálneho rozdelenia:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}}$$

Distribučná funkcia normovaného normálneho rozdelenia:

$$F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

Graf:



Laplaceov integrál – je integrál normovaného normálneho rozdelenia s premenlivou hornou hranicou, ktorého dolná hranica sa rovná 0.

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{u^2}{2}} du$$

$$F(t) = \Phi(t) + 0{,}5$$
$$t = 0.....\Phi(t) = 0....F(t) = 0{,}5$$
$$\infty.....\Phi(\infty) = 0{,}5.....F(\infty) = 1$$
$$ak...t\langle 0..... \Rightarrow -t.....\Phi(-t) = -\Phi(t).....F(-t) = 1 - F(t)$$
$$-\infty....\Phi(-\infty) = -0{,}5.......F(-\infty) = 0$$

Príklad:
Pri hromadnej výrobe výrobku je priemerná hodnota sledovaného parametra 500 mm .
Rozmer výrobkov má normálne rozdelenie so smerodajnou odchýlkou σ = 10 mm.
Aká je pravdepodobnosť, že náhodne vybraný výrobok bude mať rozmer v hraniciach (485, 515) mm.
T_o = 485 - dolná tolerancia
T_H = 515 - horná hranica
$P(485 \leq X \leq 515) = F(x_2) - F(x_2) = F(t_2) - F(t_1)$

Pre riešenie použijeme tabuľky N(O, 1),
Normálne rozdelenie s parametrami (500, 10) - pretransformujeme na N(O, 1).

$$t_1 = \frac{x_1 - \bar{x}}{\sigma} = \frac{485 - 500}{10} = -1{,}5$$

$$t_2 = \frac{x_2 - \bar{x}}{\sigma} = \frac{515 - 500}{10} = 1{,}5$$

$F(x_2) - F(x_1)$

$$F(t_2) - F(t_1) = F(1{,}5) - F(-1{,}5) = F(1{,}5) - [1 - F(1{,}5)] = \\ = 2F(1{,}5) - 1 = 2.0{,}9332 - 1 = 0{,}8664$$

$$\Phi(t_2) - \Phi(t_1) = \Phi(1{,}5) - \Phi(-1{,}5) = 0{,}8664$$

Hodnoty distribučnej funkcie resp. laplaceovho integrálu sa určia z tabuliek N(0,1)

Pravdepodobnosť, že NV bude nadobúdať hodnotu menšiu ako 485:

$$P(x \leq 485) = P(0 \leq x \leq 485) = \Phi(t_2) - \Phi(t_1) = \Phi(-1{,}5) - \Phi(-50) = \\ = -\Phi(-1{,}5) - \Phi(-50) = -\Phi(1{,}5) - [-\Phi(50)] = 0{,}5 - 0{,}4332 = 0{,}0668$$

$$t_1 = \frac{0 - 500}{10} = -50 \ ;$$

$t_2 = -1{,}5$

Obdobný postup možno použiť pre: $P(515 \leq x) = P(515 \leq x < \infty)$

Logaritmicko-normálne rozdelenie

Je asymetrickým rozdelením, je zošikmené doľava

Možno ho použiť, ak náhodná veličina X je asymetricky rozdelená, ale jej logaritmy majú normálne rozdelenie. Vytvorí sa nová NV Y
Y = log x

4.5. χ^2 (chí kvadrát) rozdelenie

Ak NV X má normálne rozdelenie, všetky hodnoty umocnené na 2 budú mať chí kvadrát rozdelenie.

$$\left.\begin{matrix} x_1 \\ x_2 \end{matrix}\right\} \quad N\,(O,\,1) \qquad\qquad \chi^2 = \sum_{i=1}^{n} x_i^2$$

$$k = n - 1$$

parameter k - počet stupňov voľnosti

Rozptyl má χ^2 rozdelenie: $\sigma^2 = \sum \left(x_i - \bar{x}\right)^2 p_i$ - pri testovaní hypotéz

4.6 T - rozdelenie (Studentovo)

Vznikne transformáciou z dvoch NV : $z = \dfrac{x}{\sqrt{\dfrac{Y}{k}}}$

NV X má normálne rozdelenie N(O, 1)
NV Y má χ^2 rozdelenie s k = n - 1 *(počet stupňov volnosti)*

používa sa pri výberových metódach, ak súbory sú malých rozsahov, čiže ak rozsah výberového súboru

$n < 30$

4.7 F - rozdelenie (Snedecorovo)

Ak NV vznikne ako podiel dvoch NV, ktoré majú χ^2 rozdelenie.
Napr. podiel dvoch rozptylov - má F-rozdelenie

rozplyt výberového súboru sa označuje s^2

$$\bar{x}_1 \ldots\ldots\ldots s_1^2$$

$$\bar{x}_2 \ldots\ldots\ldots s_2^2$$

$$\frac{s_1^2}{s_2^2} = F$$

Parametre F-rozdelenia sú k_1 a k_2:

$k_1 = n_1 - 1$ n_1 je rozsah jedného výberového súboru
$k_2 = n_2 - 1$ n_2 je rozsah druhého výberového súboru

Aproximácia (vyrovnanie, nahradenie) empirického rozdelenia vhodným typom teoretického rozdelenia

Aproximácia spočíva v nasledovných krokoch:

1. výber modelu vhodného teoretického rozdelenia
2. určenie parametrov teoretického rozdelenia z empiricky získaných údajov (empirického súboru)
3. výpočet teoretických početností.

Pr.

Interval	n_i	p_i	t_1	t_2	$\Phi(t_2)$	$\Phi(t_1)$	p_2'	n'_i	zaokr úhl
-13,25>	**20**	**0,04**		**-1,875**	**-0,468**	**-0,5**	**0,032**	**15,9**	**16**
(13,25-13,30>	**35**	**0,07**	**-1,875**	**-1,25**	**-0,3944**	**-0,468**	**0,0738**	**36,9**	**37**
(13,30-13,35>	**80**	**0,16**	**-1,25**	**-0,625**	**-0,24**	**-0,3944**	**0,1544**	**77,2**	**77**
(13,35-13,40>	**105**	**0,21**	**-0,625**	**0**	**0**	**-0,24**	**0,24**	**120**	**120**
(13,40-13,45>	**140**	**0,28**	**0**	**0,625**	**0,24**	**0**	**0,24**	**120**	**120**
(13,45-13,50>	**60**	**0,12**	**0,625**	**1,25**	**0,3944**	**0,24**	**0,1544**	**77,2**	**78**
(13,50-13,55>	**35**	**0,07**	**1,25**	**1,875**	**0,468**	**0,3944**	**0,0738**	**36,9**	**37**
(13,55-	**25**	**0,05**	**1,875**	**∞**	**0,5**	**0,468**	**0,032**	**15,9**	**16**
	500	**1**							**500**

1. Ako model bolo zvolené normálne rozdelenie: N(x, σ)

2. Výpočet parametrov normálneho rozdelenia

$$\bar{x} = \upsilon_1 = \frac{\sum x_i n_i}{n}$$

$$\upsilon = \frac{\sum x_i^2 n_i}{n}$$

$$\sigma = \sqrt{\upsilon_2 - \upsilon_1^2}$$

x = 13,4, σ = 0,08 - výsledok

n_i' - teoretické početnosti

$$p_i = \frac{n_i}{n}$$

$$n'_i = p'_i \,.n$$

3. Výpočet teoretických pravdepodobností

p'_i - teoretické pravdepodobnosti

$$P\left(x_1 \le x \le x_2\right) = \Phi(t_2) - \Phi(t_1) = p'_i$$

$$t_1 = \frac{x_1 - \overline{x}}{\sigma}$$

$$t_2 = \frac{x_2 - \overline{x}}{\sigma} = \frac{13{,}25 - 13{,}4}{0{,}08} = -1{,}875$$