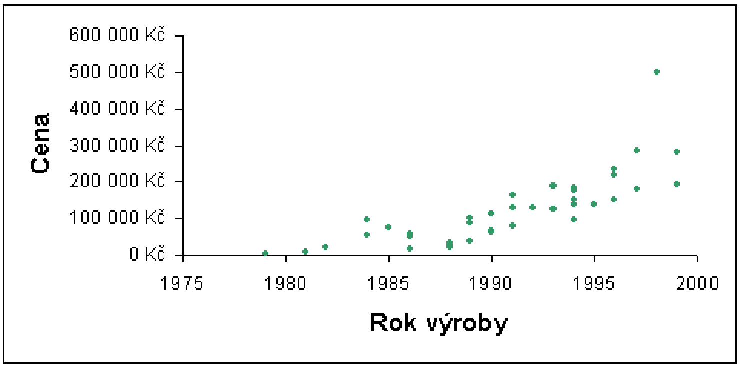
Sedem základných nástrojov manažérstva kvality v doprave a logistike - bodový diagram

Bodový diagram (angl. Scatter Plots, Scatterplot alebo scattergraph) predstavuje grafickú metódu, ktorou zachytávame vzájomný vzťah medzi dvomi premennými. Pomocou bodového diagramu možno posudzovať napríklad vzájomnú súvislosť medzi dvomi znakmi kvality výrobku, súvislosti medzi určitým znakom kvality výrobku a jednotlivými parametrami procesu a pod.

Bodový diagram používame, ak chceme kontrolovať nejakú premennú. Riadiaci parameter alebo nezávislá premenná sa zvyčajne vynesie na vodorovnú os. Meraná alebo závislá premenná sa zvyčajne zobrazuje na zvislej osi. Vynášaním jednotlivých meraní sa ilustruje len stupeň korelácie – nie príčinné súvislosti medzi dvoma premennými. Dáta sa pritom zobrazujú ako súbor bodov v karteziánskej sústave súradníc.

Uvedený diagram sa v doprave veľmi často používa. Na obr. 1 napr. vidíme graf zachytávajúci vplyv roku výroby na cenu predávaného automobilu.



Obr. 1 Závislosť ceny vozidla od roku jeho výroby [http://www.koppel.cz/VSE/files/STP202.htm]

Bodový diagram môže zachytávať rôzne stupne či druhy korelácií medzi dvoma premennými s určitým intervalom spoľahlivosti. Korelácia pritom môže byť pozitívna – rast alebo negatívna – pokles príp. nulová, teda žiadna. Pre vyhodnocovanie kvality je pritom dôležité, či sa vynesené „bodky" snažia dodržiavať určitú líniu – vyskytujú sa v okolí nejakej pomyselnej priamky príp. krivky, ktorá by mohla reprezentovať nelineárne vzťahy medzi premennými.

Pri analýze korelácie využívame líniovú závislosť, teda pozorujeme, či sa vynesené body približujú k fiktívnej priamke, tzv. trendovej čiare. Pri preložení súboru bodov trendovou čiarou by sme mali mať na pamäti, že toľko bodov, koľko leží pod čiarou, by malo ležať aj pod čiarou. V tomto prípade môžeme využiť rovnicu priamky y = a. x pre odhad hodnôt, ktoré sú mimo nášho súboru. Pre lepší odhad charakteru korelácie je možné skupinou bodov s využitím metódy najmenších štvorcov teda preložiť priamku a tak určiť tzv. Pearsonov koeficient korelácie „r".

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{(x_i^2 - n\bar{x}^2)(y_i^2 - n\bar{y}^2)}} \tag{1}$$

kde:

$\bar{x}$........ aritmetický priemer hodnôt premennej x, vypočítaný vzťahom:

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

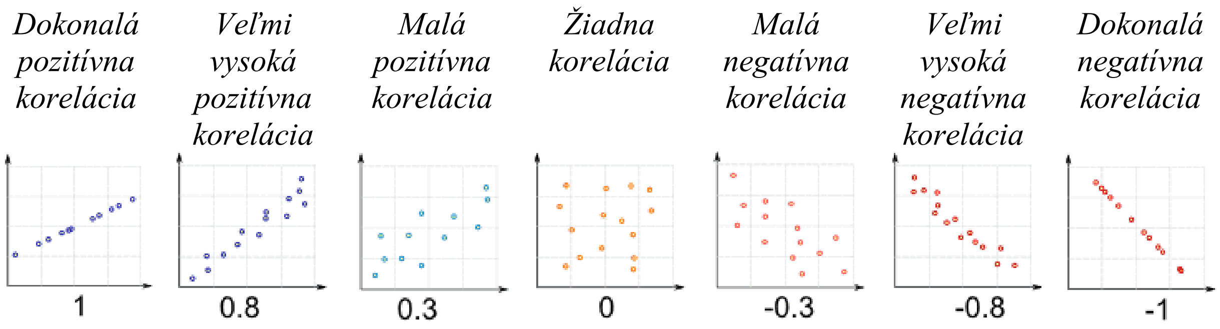
$\bar{y}$........ aritmetický priemer hodnôt premennej y, vypočítaný vzťahom:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \quad (3)$$

Pearsonov koeficient korelácie môže nadobúdať hodnoty z intervalu od -1 po 1. Koreláciu potom možno interpretovať v závislosti od výsledku r:

- ak r = 0, tak korelácia je *nulová*;
- ak r = < 0,1, korelácia je *triviálna*;
- ak r = (0,1÷ 0,3), korelácia je *malá*;
- ak r = (0,3 ÷ 0,5), korelácia je *stredná*;
- ak r = (0,5 ÷ 0,7), korelácia je *vysoká*;
- ak r = (0,7 ÷ 0,9), korelácia je *veľmi vysoká*;
- ak r = (0,9 ÷ 0,999..), korelácia je *takmer dokonalá*;
- ak r = 1, korelácia je *dokonalá.*

Obr. 2 zachytáva grafické zobrazenia jednotlivých situácií s určením hodnoty Pearsonovho koeficientu.



Obr. 2 Hodnoty Pearsonovho koeficientu

Treba však mať na pamäti aj skutočnosť, že niekedy vzájomná korelácia dvoch premenných nemusí znamenať ich príčinnú súvislosť, teda nehovorí o príčine či dôsledku. Korelujúce premenné môžu závisieť od nejakej inej veličiny.

Aplikáciu bodového diagramu si vysvetlíme na nasledujúcom príklade.

Príklad 1:

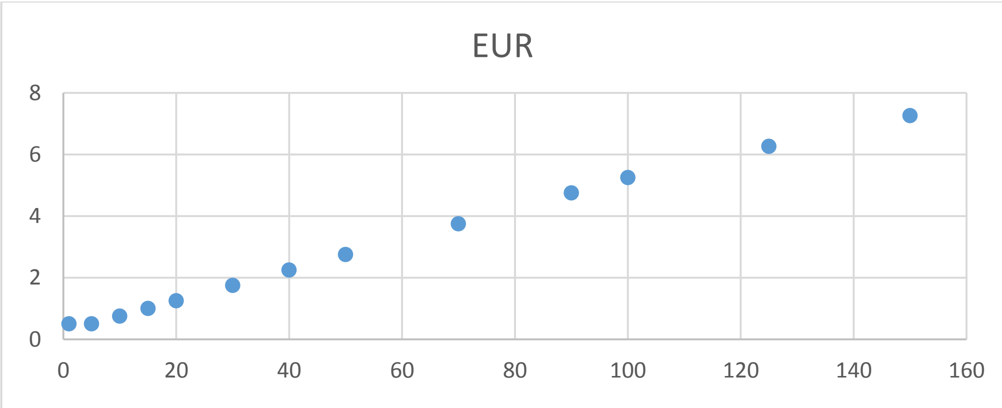
Pri skúmaní cenníka používaného v reálnej taxislužbe chceme zistiť, aká je závislosť medzi prepravnou vzdialenosťou a cenou cestovného. Ak existuje lineárna závislosť, nemá zmysel deliť si celkovú vzdialenosť na kratšie úseky.

Z cenníka taxislužby sa dozvedáme nasledovné údaje (tabuľka 1):

Tabuľka 1 Cena prepravného v reálnej taxislužbe

Závislosť ceny cestovného a prepravnej vzdialenosti													
km	1	5	10	15	20	30	40	50	70	90	100	125	150
EUR	0,5	0,5	0,75	1,00	1,25	1,75	2,25	2,75	3,75	4,75	5,25	6,26	7,26

Jednotlivé body znázorníme do grafu (obr. 3).



Obr. 3 Závislosť ceny cestovného a vzdialenosti v taxislužbe

Už z obrázku vidíme, že jednotlivé body sa „držia" určitej línie, teda medzi hodnotami existuje vysoká závislosť.

Tento predpoklad potvrdí výpočet Pearsonovho koeficientu podľa vzťahu (1). Aby sme mohli do vzťahu dosadiť správne hodnoty, je vhodné použiť pomocnú tabuľku výpočtov (tabuľka 2). Najskôr si ale vypočítame priemery jednotlivých veličín pomocou vzťahov (2) a (3) a ich druhé mocniny.

Priemerná vzdialenosť:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1+5+10+15+20+30+40+50+70+90+100+125+150}{13} = \frac{706}{13} = 54{,}3076923\ km$$

$$\bar{x}^2 = 54{,}3076923^2 = 2949{,}3254$$

Priemerná cena:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{0{,}5+0{,}5+0{,}75+1+1{,}25+1{,}75+2{,}25+2{,}75+3{,}75+4{,}75+5{,}25+6{,}26+7{,}26}{13} = \frac{38{,}02}{13} = 2{,}92461538\ EUR$$

$$\bar{y}^2 = 2{,}92461538^2 = 8{,}5533751$$

Tabuľka 2 Pomocné výpočty potrebných hodnôt

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**
x	1	5	10	15	20	30	40
y	0,5	0,5	0,75	1,00	1,25	1,75	2,25
x.y	0,5	2,5	7,5	15	25	52,5	90
x^2	1	25	100	225	400	900	1600
y^2	0,25	0,25	0,5625	1	1,5625	3,0625	5,0625

Pokračovanie tabuľky 2

n	8	9	10	11	12	13	***Spolu***
x	50	70	90	100	125	150	*706*
y	2,75	3,75	4,75	5,25	6,26	7,26	*38,02*
x.y	137,5	262,5	427,5	525	782,5	1089	*3417*
x^2	2500	4900	8100	10000	15625	22500	*66876*
y^2	7,5625	14,0625	22,5625	27,5625	39,1876	52,7076	*175,3952*

A výpočet Pearsonovho koeficientu:

$$r = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{(x_i^2 - n\bar{x}^2)(y_i^2 - n\bar{y}^2)}} = \frac{3417 - 13.54{,}3076923.2{,}92461538}{\sqrt{(66876 - 13.2949{,}3254)(175{,}3952 - 13.8{,}5533751)}} = 0{,}999053$$

Výpočet potvrdzuje, že medzi prepravnou vzdialenosťou a cenou za túto prepravu existuje silná závislosť, ktorá je takmer lineárna.

Ako už bolo spomínané v texte, pri výpočtoch a odhadoch je možné využiť mieru vzájomnej závislosti skúmaných dvoch veličín a pomocou preloženej trendovej priamky zisťovať hodnoty, ktoré v základnom skúmanom súbore neboli.

Pri preložení súboru bodov trendovou čiarou (angl. Line of Best Fit) platí pravidlo, že by mala byť tak blízko všetkým bodom, ako je len možné a ten istý počet bodov má byť nad ňou, ako je pod ňou.

Po preložení súboru bodov trendovou čiarou môžeme odhadovať ďalšie hodnoty využitím interpolácie a extrapolácie.

Interpoláciou hľadáme hodnoty medzi našimi dátami. Extrapoláciou hľadáme hodnoty za hranicou našich dát. Uvedené metódy v prípade tarifného cenníka nemožno aplikovať, no možno ich odporúčať na aplikáciu pri zjednávaní ceny. Pripomínam, že medzi hodnotami súboru musí existovať silná závislosť. Vtedy môžeme pre trendovú čiaru určiť tzv. sklon priamky „m“:

$$m = \frac{zmena\ v\ y}{zmena\ v\ x} \tag{4}$$

Uveďme si príklad presnejšieho výpočtu:

Príklad 2:

O akej cene by sa dalo uvažovať, ak by sme použili uvedený cenník taxislužby, no cestovali by sme 11 km a 170 km?

Najskôr si musíme určiť sklon m. Ide o pomer zmien hodnôt v dvoch nameraných súboroch. Vyberieme si dve hodnoty (5: 0,5 a 125:6,26). Použijeme vzťah (4).

$$m = \frac{zmena\ v\ y}{zmena\ v\ x} = \frac{6{,}26 - 0{,}5}{125 - 5} = 0{,}048$$

Teraz „sklon“ zahrnieme do vzťahu:

$y - y_1 = m(x - x_1)$, teda:

$$y = m(x - x_1) + y_1 \tag{5}$$

Takto sme si určili rovnicu našej trendovej priamky, pomocou ktorej budeme interpolovať a extrapolovať zisťované hodnoty. Pri výpočte si pomôžeme najbližšími nameranými hodnotami.

Pri jazde taxíkom na vzdialenosť 11 km si vyberieme najbližšiu hodnotu z tarify 10 km, za ktorú zaplatíme 0,75 €. Použijeme vzťah (5).

$y = m(x - x_1) + y_1$
$y = 0{,}048(11 - 10) + 0{,}75 = 0{,}798$ t. j. 0,80 €

Pri jazde taxíkom na vzdialenosť 170 km si vyberieme najbližšiu hodnotu z tarify 150 km, za ktorú zaplatíme 7,26 €. Použijeme vzťah (5).

$y = m(x - x_1) + y_1$
$y = 0{,}048(170 - 150) + 7{,}26 = 8{,}22$ €

Uvedené výpočty vytvárajú rámec, do akého možno cenu zjednávať.

Hodnoty sú blízke tým, ktoré máme uvedené v grafe. Nie sú však presné, ale približné. Upozorňujem, že pri používaní podobných prepočtov nie je vhodné extrapolovať príliš vysoké hodnoty (Koľko zaplatím taxíkom za cestu 1234 km?) alebo príliš nízke hodnoty, ktoré nie sú reálne (Koľko zaplatím taxíkom za cestu -50 km?). Pri použití záporných hodnôt dokonca získame záporné hodnoty platieb, čo je nemožné.´

Článok poukazuje na možnosť použitia korelačného diagramu pri určovaní ceny za prepravu.

Vlastná analýza tohto diagramu môže poskytnúť cenné prvotné informácie a študovanom znaku kvality, resp. o procese, ktorého je tento znak kvality výsledkom. V prípade korelačného diagramu nerozoznávame nezávislú a závislú premennú, ale obe považujeme za náhodné premenné. Vzájomne si zodpovedajúce hodnoty dvoch premenných vynášame ako body v karteziánskej súradnicovej sústave. Podľa tvaru skupiny bodov potom môžeme usudzovať o možnej korelácii týchto premenných. Ak body zľava doprava stúpajú, hovoríme o pozitívnej korelácii, ak klesajú, o negatívnej korelácii. Ak tvar nemá viditeľnú orientáciu, veličiny pravdepodobne nekorelujú alebo korelujú nelineárne. [1]

Literatúra:

[1] http://www.koppel.cz/VSE/files/STP202.htm
[2] http://www.podbrezovan.sk/regionnoviny/podbrezovan.nsf/page/2014_22_Sedem_zakladnych_nastrojov_manazerstva_kvality_I_
[3] http://portal2.tuke.sk/hf-kim/bakalar/predmety-bc/integrovane-manazerske-systemy/podklady-cvicenia/IMS_10._cast.pdf

Ing. Adela Poliaková, EUR ING, PhD.
Slovenská technická univerzita v Bratislave
Materiálovotechnologická fakulta so sídlom v Trnave
Ústav bezpečnosti, environmentu a kvality