

Ústav lékařské informatiky, 2. LF UK

STATISTIKA III.

2005

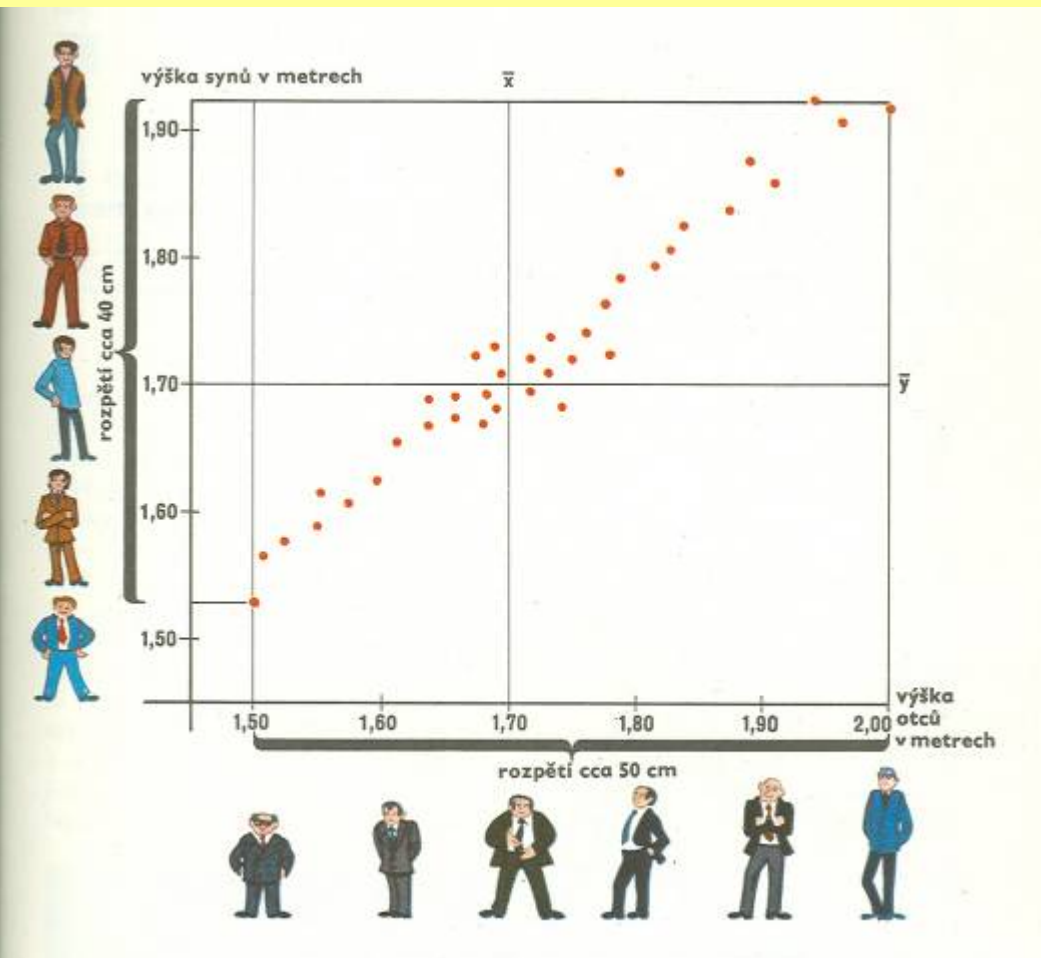


PŘEHLED TÉMATIKY

1. Regrese
2. Časové řady
3. Korelace
4. Vícerozměrné metody

REGRESE

ZÁKLADNÍ POJMY



Nezávisle proměnná –
výchozí, ovlivňující znak
(kupř. věk dítěte)

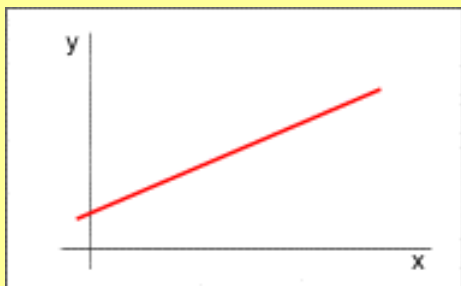
Závisle proměnná – odvozený,
ovlivňovaný znak (kupř.
výška dítěte)

Regresní rovnice – popisuje
typ závislosti

Regresní odhad – dosazením
do rovnice lze odhadnout
hodnotu závisle proměnné

Regrese

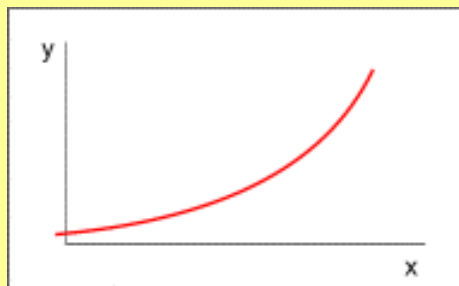
Matematický vztah mezi dvěma závislými znaky



$$y = a + b * x$$

(př. viz níže)

$$\text{výška} = 80 \text{ cm} + 5 * \text{věk}$$



$$y = a * e^{bx}$$

(př. zátěž - TF)

$$\text{váha} = 8 \text{ kg} + 2 * \text{věk}$$



$$y = a * \ln(b * x)$$

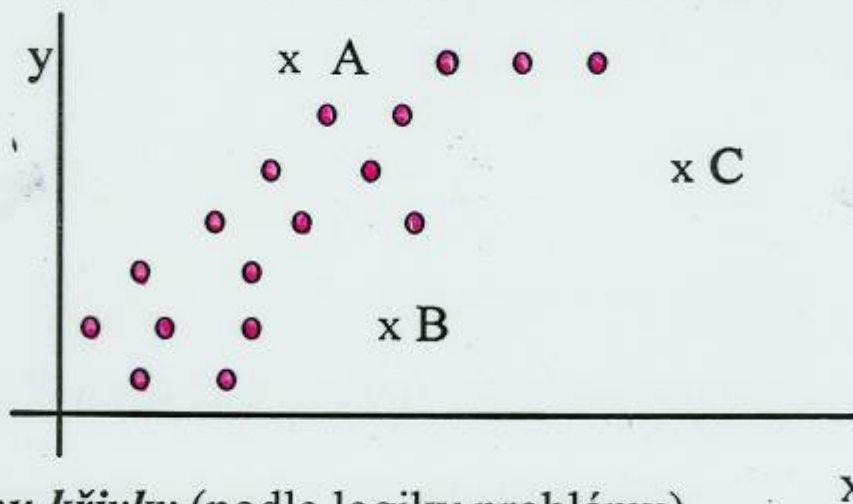
(př. Saturace Hb – O₂)



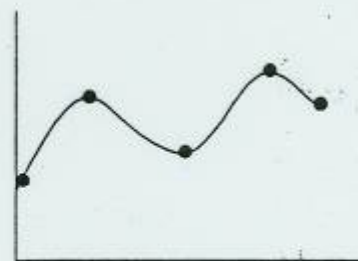
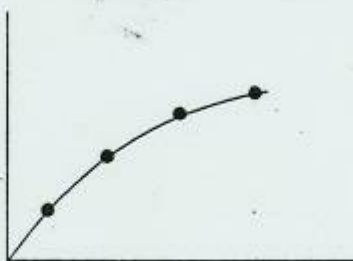
Příklady rovnic užívaných v pediatrické praxi pro odhad přiměřené výšky a váhy v předškolním věku, kdy je závislost zhruba lineární

REGRESE – PRACOVNÍ POSTUP I.

1. *Konstrukce bodového grafu* (problém odlehých hodnot)

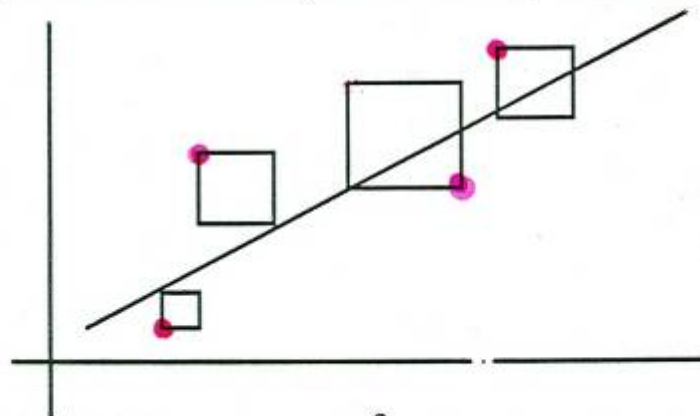


2. *Volba typu křivky* (podle logiky problému)

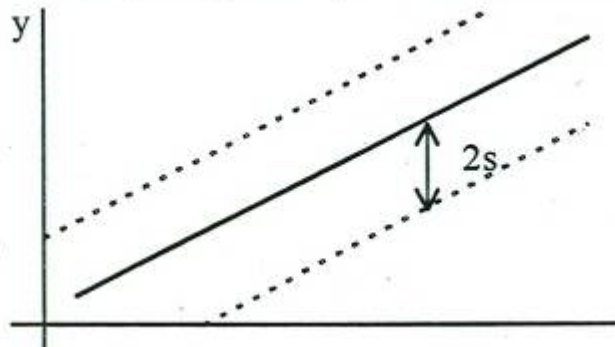


REGRESE – PRACOVNÍ POSTUP II.

3. *Výpočet regresní rovnice* (metoda nejmenších čtverců)

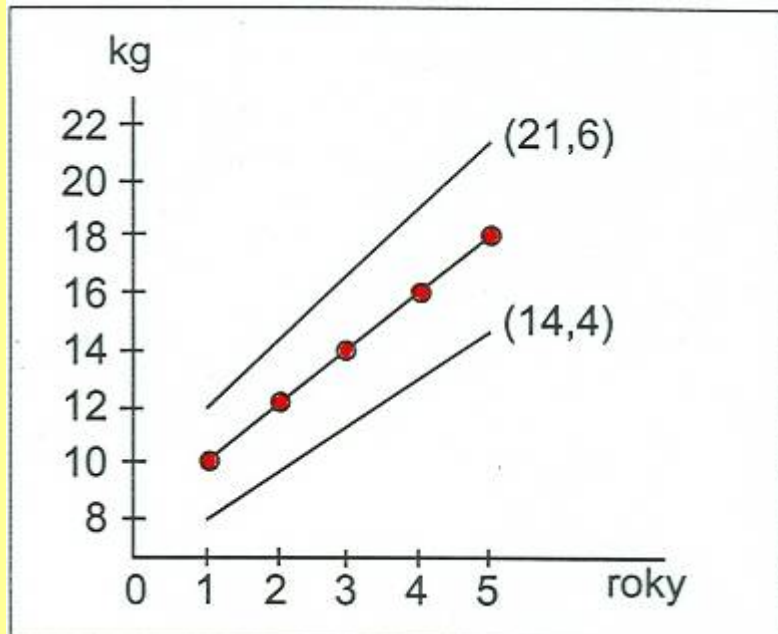


4. *Hodnocení kvality regrese* (S^2 - chyba odhadu)



REGRESE VYUŽITÍ

Regresní odhad

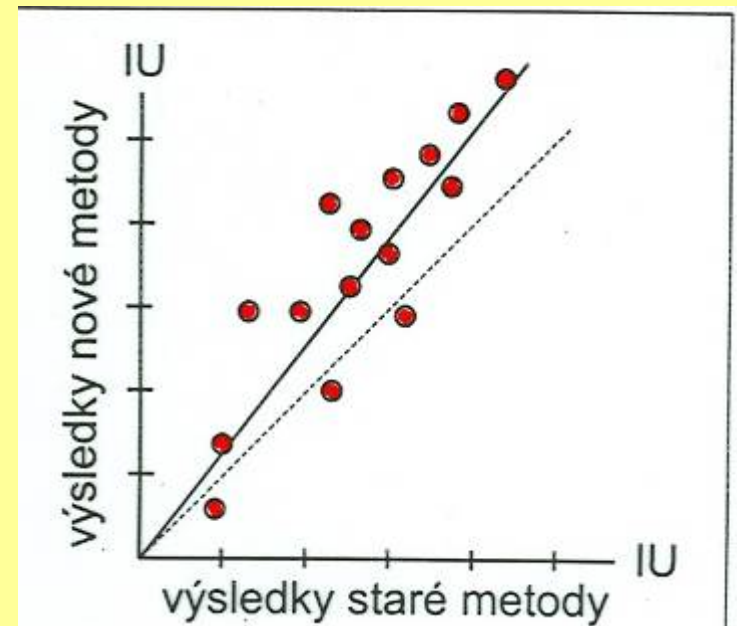


Posuzování výběru

$$Y = AX + B$$

$$\text{výpočet } Y = 0,1 X + 0,2$$

Hypotéza a regresní vztah



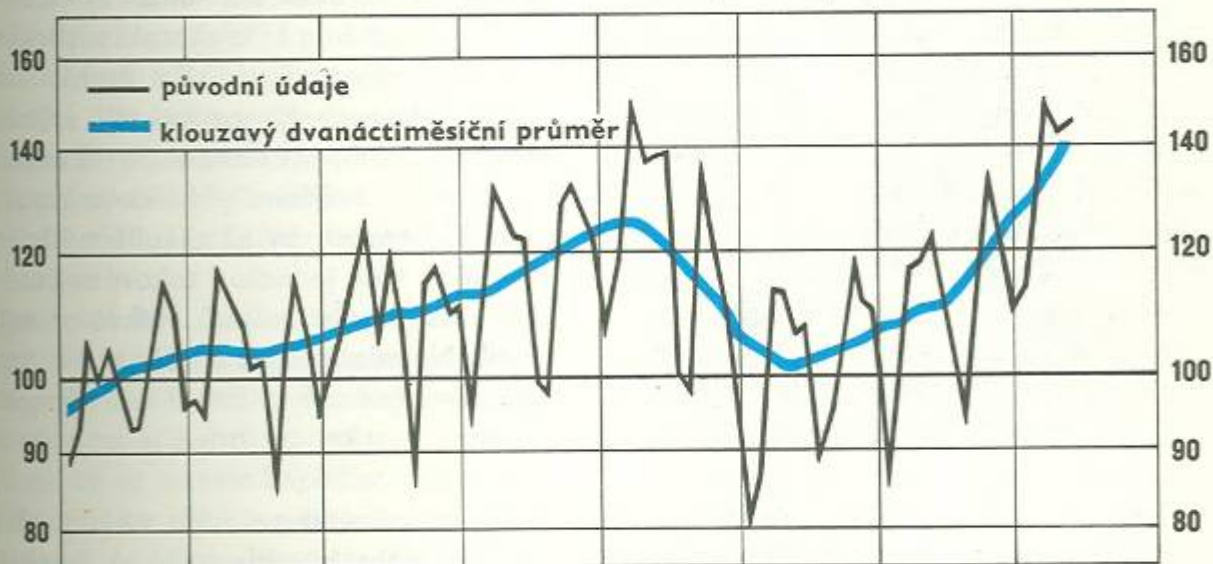
Posuzování výběru

$$Y = AX + B$$

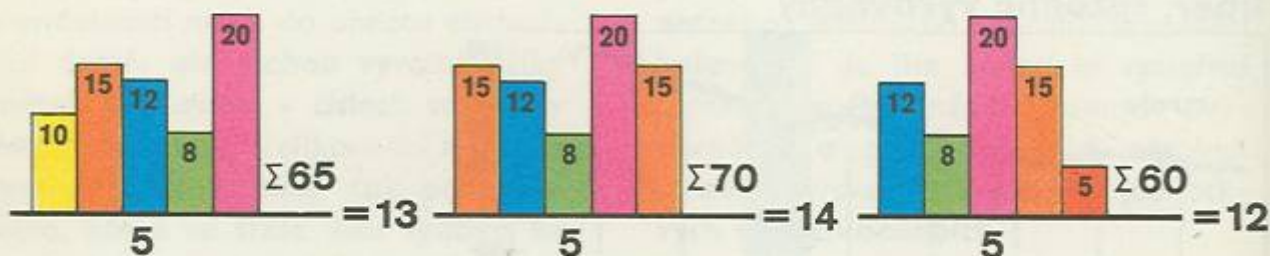
$$\text{výpočet } Y = 0,1 X + 0,2$$

ČASOVÉ ŘADY – KLOUZAVÝ PRŮMĚR

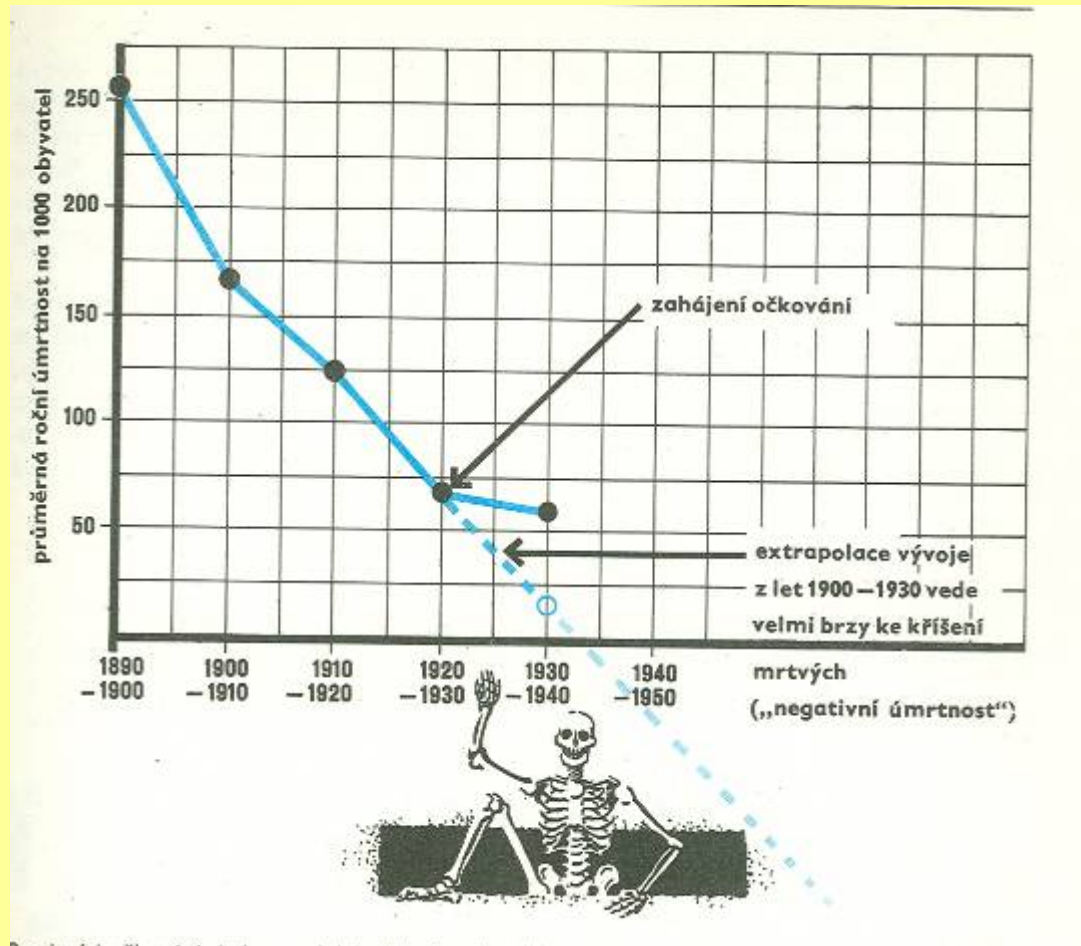
Vnitrostátní obrát v automobilovém průmyslu 1962 = 100



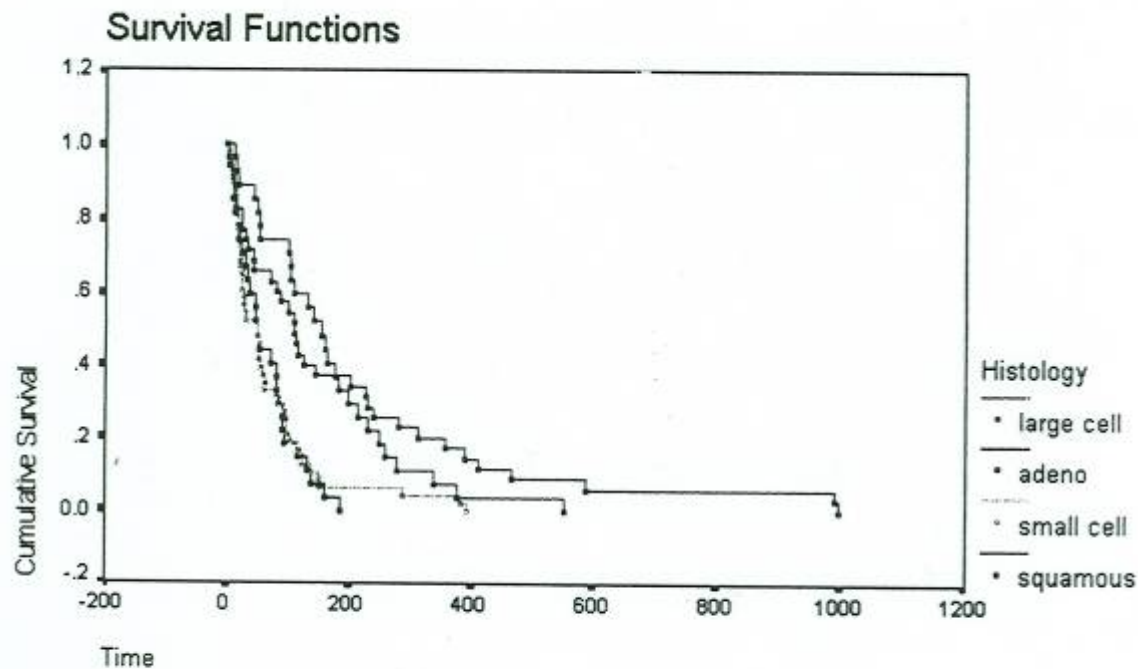
„Klouzavý“ průměr také v grafickém znázornění vykazuje kolísavou křivku. Hektický sestup a vzestup během roku se uklidňují a dostáváme klidnější křivku vývoje. Čísla jsou za léta 1962



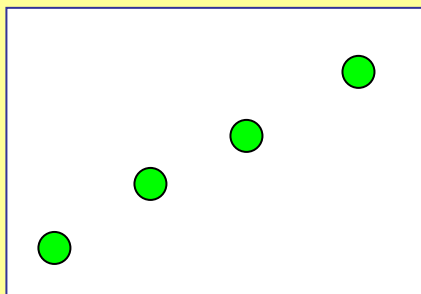
ČASOVÉ ŘADY - EXTRAPOLACE



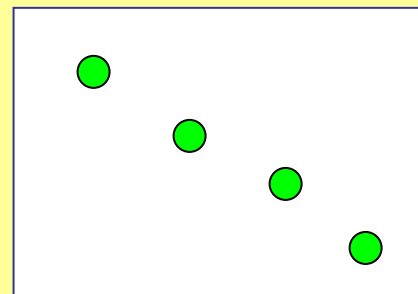
ČASOVÉ ŘADY – KŘIVKY PŘEŽITÍ



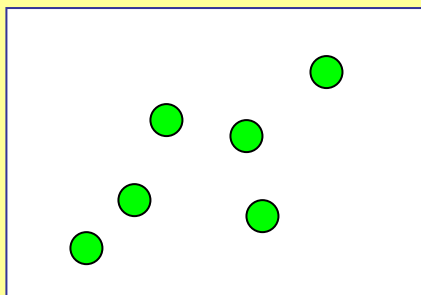
KORELACE - ZÁKLADNÍ POJMY



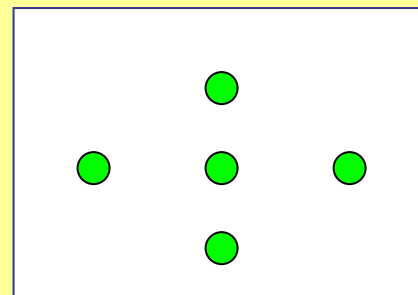
Přímá úměra $r = +1$



Nepřímá úměra $r = -1$



***Příklad částečné
závislosti $r = 0,3$***



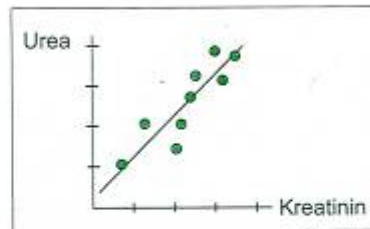
Nezávislost $r = 0$

KORELACE - VYUŽITÍ

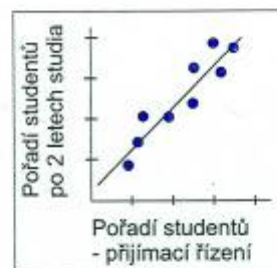
(Pearsonův) korelační koeficient – pro metrická data

$$r = \frac{\sum x - x_i \cdot \sum y - y_i / n - 1}{S_x \cdot S_y}$$

Př. vztah urea – kreatinin



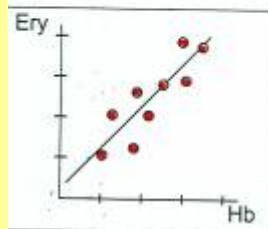
Spearmanův koeficient – pro ordinální data (korelace 2 pořadí)



Test významnosti

$H_0: r = 0$

Je vztah mezi Hb a Ery významný když při hodnocení 15 osob $r = 0,7$?



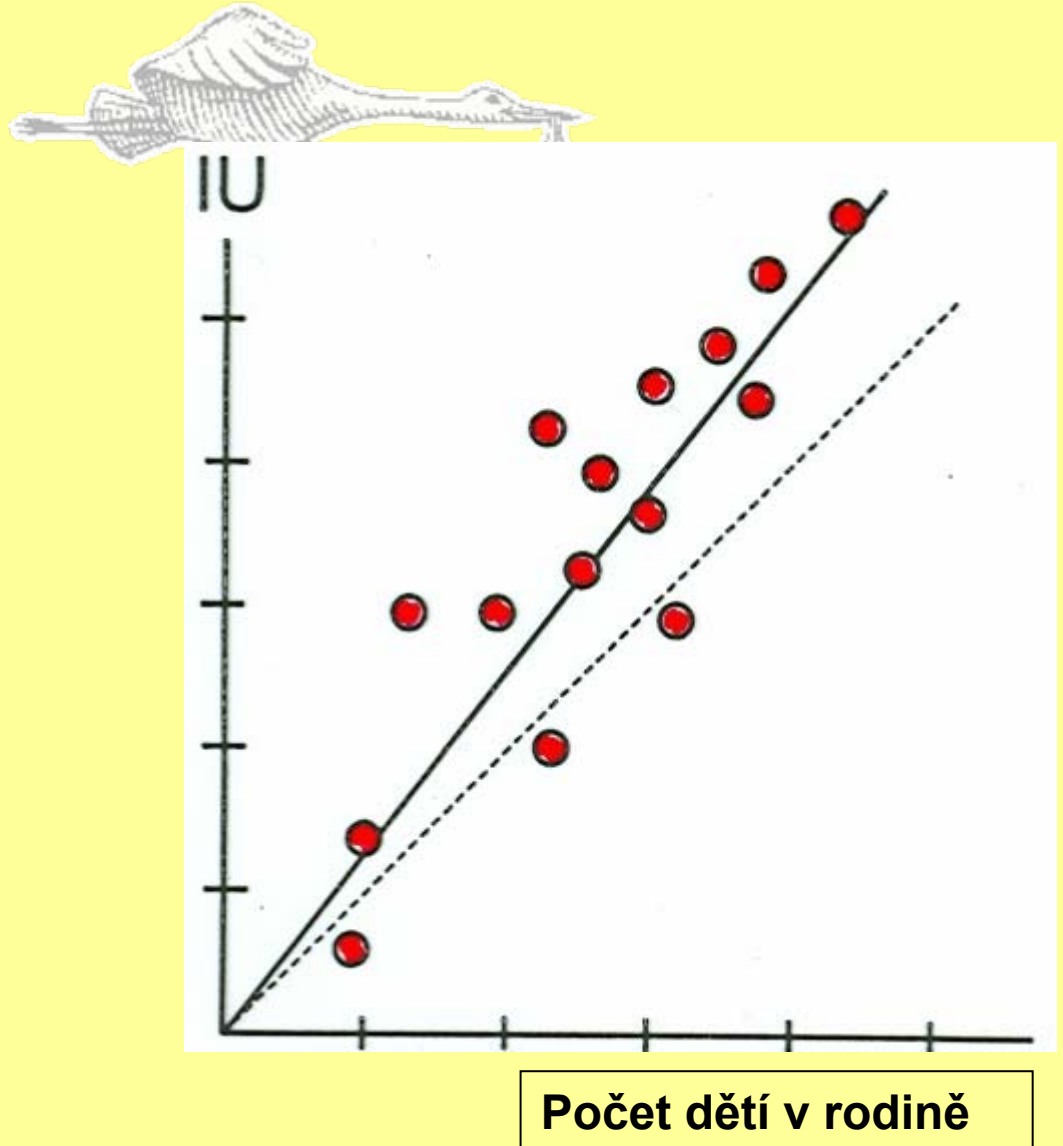
Statistické tabulky

n	Hodnota r pro $\alpha = 0,05$
5	0,75
10	0,57
15	0,48
20	0,42

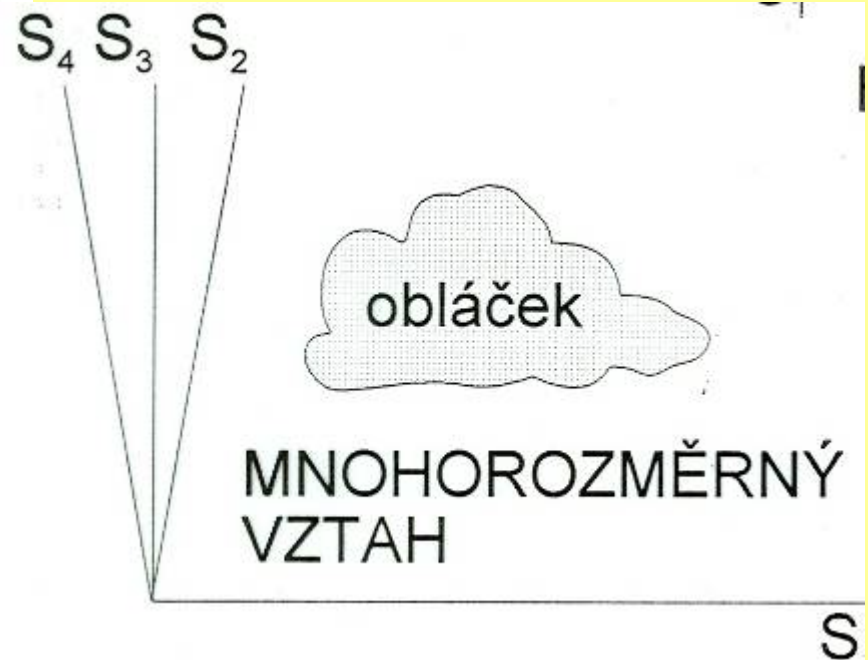
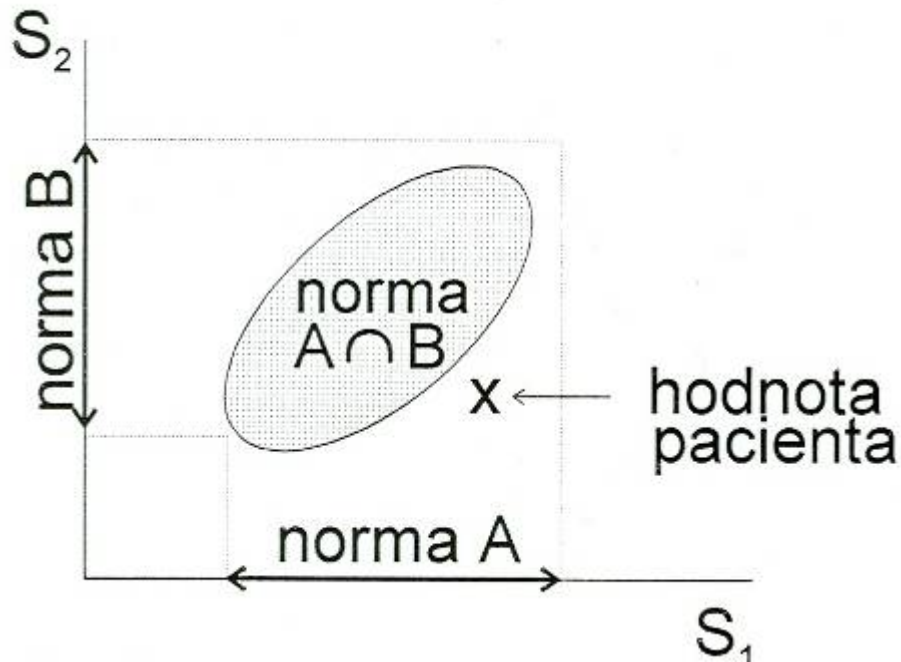
KORELACE NOMINÁLNÍCH HODNOT

Ženich	Nevěsta		
	svobodná	ovdovělá	rozvedená
svobodný	$53^2 = \frac{2809}{747} = 3,8$	$(-13)^2 = \frac{169}{20} = 8,5$	$(-40)^2 = \frac{1600}{82} = 19,5$
ovdovělý	$(-25)^2 = \frac{625}{33} = 18,9$	$9^2 = \frac{81}{1} = 81,0$	$6^2 = \frac{36}{4} = 9,0$
rozvedený	$(-38)^2 = \frac{1444}{99} = 14,5$	$4^2 = \frac{16}{3} = 5,3$	$34^2 = \frac{1156}{11} = 105,1$
			$37,2 \quad + \quad 94,8 \quad + \quad 133,6 = 265,6 = \chi^2$

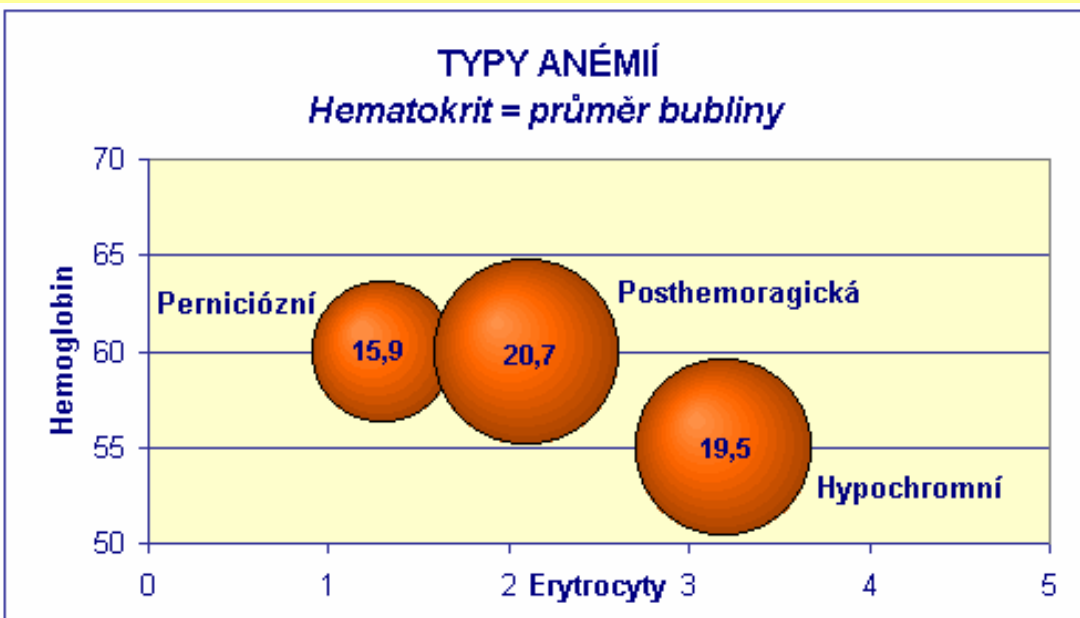
ZDÁNlivÉ KORELACE



VÍCEROZMĚRNÝ VZTAH - ZÁKLADNÍ POJMY



VÍCEROZMĚRNÝ VZTAH GRAFICKÁ REPREZENTACE



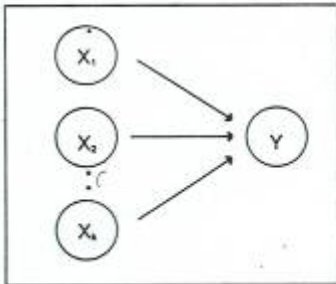
**Grafické vyjádření závislosti
3 faktorů – bublinový graf
3. faktor – velikost bublin**



VÍCEROZMĚRNÁ REGRESE

Mnohorozměrná regrese

Příklad: odhad hodnoty těžko dostupného příznaku Y ze známých příznaků



$$Y = X_1 * A + X_2 * B \dots$$

A, B ... = váhy příznaků

NOVOROZENECKÁ
ŽLOUTENKA
GESTAČNÍ VĚK
BILIRUBIN
GESTAČNÍ HMOTNOST
VOLNÉ MASTNÉ
KYSELINY
CELKOVÁ BÍLKOVINA



LÉČBA

Vazebná
kapacita
albuminu
(VKA)

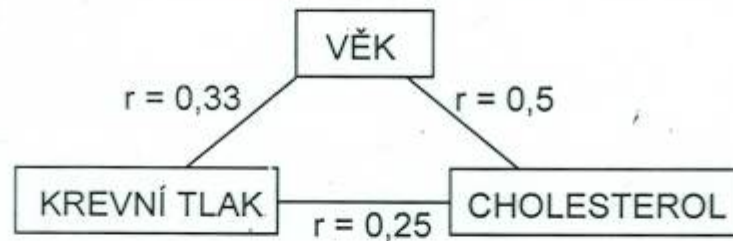
↓ VKA	OBSERVACE
Stř. VKA	FOTOTERAPIE
↑ VKA	VÝMĚNNÁ TRANSFÚZE

KORELACE VÍCE FAKTORŮ

Mnohonásobná korelace

Posuzování všech faktorů najednou

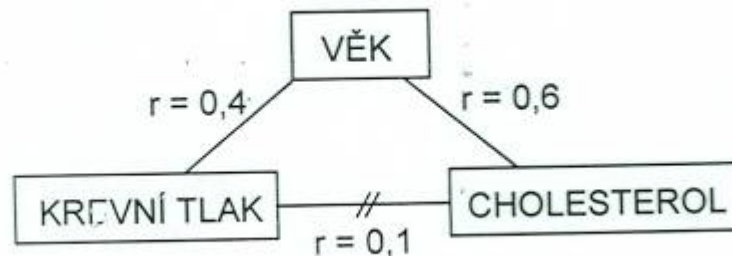
Odpověď na otázku – „Je pacient při daném věku, tlaku a cholesterolu v oblasti normálu?“



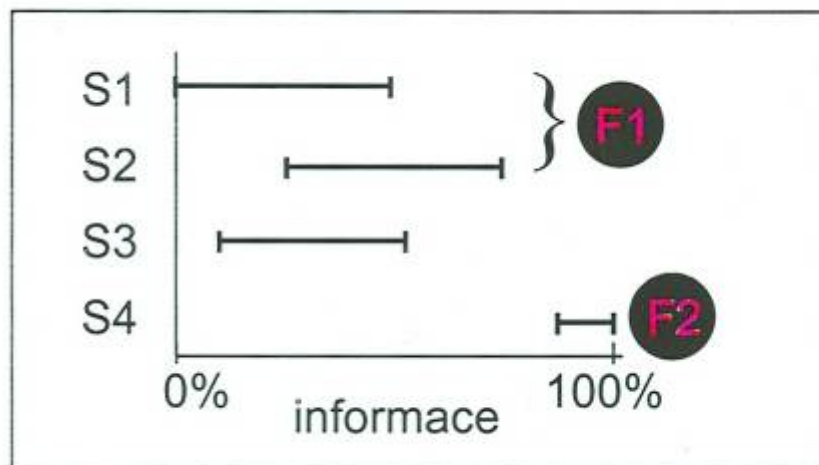
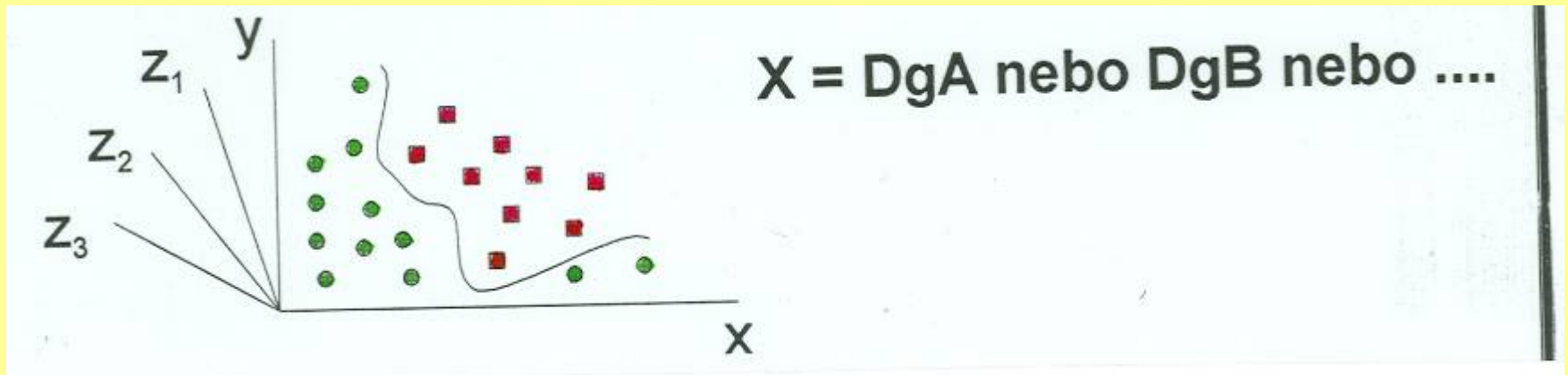
Parciální korelace

Ověření vlivu dalšího faktoru na studovaný vztah

Odpověď pro otázku „Které faktory jsou v příčinném vztahu?“



DISKRIMINAČNÍ ANALÝZA



Faktory:

$$F1 = (S1 + S2)/2$$

$$F2 = S4$$

SHLUKOVÁ ANALÝZA

$X = Dg A1 \text{ v } Dg A2 \dots\dots$

1. Shluky dle objektů

2. Shluky dle vlastností

