# ČASOVÉ ŘADY

Mgr. et Mgr. Jiří Kalina, PhD.

prof. Ing. Jiří Holčík, CSc.

UKB, pavilon D29 (Recetox), kancelář 123
kalina@mail.muni.cz

# OPERACE SE DVĚMA POSLOUPNOSTMI

## (BINÁRNÍ OPERACE)

✸

# POKRAČOVÁNÍ

# KORELACE

# DEFINICE

ABZ slovník cizích slov

**Korelace =** vzájemný vztah, souvztažnost mezi znaky, veličinami, ději

# DEFINICE

**Korelace** (z lat.) znamená vzájemný vztah mezi dvěma procesy nebo veličinami.

# DEFINICE

**Korelace** (z lat.) znamená vzájemný vztah mezi dvěma procesy nebo veličinami.

lātiō - nesení, poskytování

relātiō – nesení zpět, odnášení, opakování; zpráva; vztah, poměr

correlātiō – vzájemný vztah, souvislost

# DEFINICE

**Korelace** (z lat.) znamená vzájemný vztah mezi dvěma procesy nebo veličinami.

Pokud se jedna z nich mění, mění se **korelativně** i druhá a naopak.

Pokud se mezi dvěma procesy ukáže korelace, je možné, že na sobě závisejí, nelze z toho však ještě usoudit, že by jeden z nich musel být příčinou a druhý následkem. To samotná korelace nedovoluje rozhodnout.

# DEFINICE

**Korelace** (z lat.) znamená vzájemný vztah mezi dvěma procesy nebo veličinami.

Pokud se jedna z nich mění, mění se **korelativně** i druhá a naopak.

**Kauzalita** - příčinná souvislost či závislost. Jeden jev vyvolává druhý, popřípadě se oba vzájemně podporují (synergie)

# DEFINICE

Ve specifičtějším slova smyslu se pojem korelace užívá ve **statistice**, kde znamená vzájemný (nejčastěji **lineární**) vztah mezi znaky či veličinami ***x*** a ***y***.

# DEFINICE

Ve specifičtějším slova smyslu se pojem korelace užívá ve **statistice**, kde znamená vzájemný (nejčastěji **lineární**) vztah mezi znaky či veličinami ***x*** a ***y***.

funkční závislost a míru této závislosti

## KORELAČNÍ KOEFICIENT

numerická míra určitého typu korelace, tj. statistického vztahu mezi dvěma proměnnými

# ČASOVĚ NEZÁVISLÁ DATA

## KORELAČNÍ KOEFICIENT

1. $$\rho_{X,Y} = \frac{\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

2. $$r_s = \rho_{\mathrm{rg}_X,\mathrm{rg}_Y} = \frac{\mathrm{cov}(\mathrm{rg}_X,\mathrm{rg}_Y)}{\sigma_{\mathrm{rg}_X}\sigma_{\mathrm{rg}_Y}}$$

3. $$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

4. $$\tau = \frac{(\text{počet shodných dvojic}) - (\text{počet neshodných dvojic})}{n(n-1)/2}$$

# ČASOVĚ NEZÁVISLÁ DATA

## KORELAČNÍ KOEFICIENT

1. $$\rho_{X,Y} = \frac{\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

**obecný vztah**

### PEARSONŮV KORELAČNÍ KOEFICIENT

3. $$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

**výběrový Pearsonův korelační koeficient**

# ČASOVĚ NEZÁVISLÁ DATA

## KORELAČNÍ KOEFICIENT

**pořadová korelace** – míra vztahu mezi pořadím hodnot dvou proměnných

**2.** $$r_s = \rho_{\mathrm{rg}_X, \mathrm{rg}_Y} = \frac{\mathrm{cov}(\mathrm{rg}_X, \mathrm{rg}_Y)}{\sigma_{\mathrm{rg}_X} \sigma_{\mathrm{rg}_Y}}$$

**SPEARMANŮV POŘADOVÝ KORELAČNÍ KOEFICIENT**

**KENDALLŮV POŘADOVÝ KORELAČNÍ KOEFICIENT**

**4.**

$$\tau = \frac{(\text{počet shodných dvojic}) - (\text{počet neshodných dvojic})}{n(n-1)/2}$$

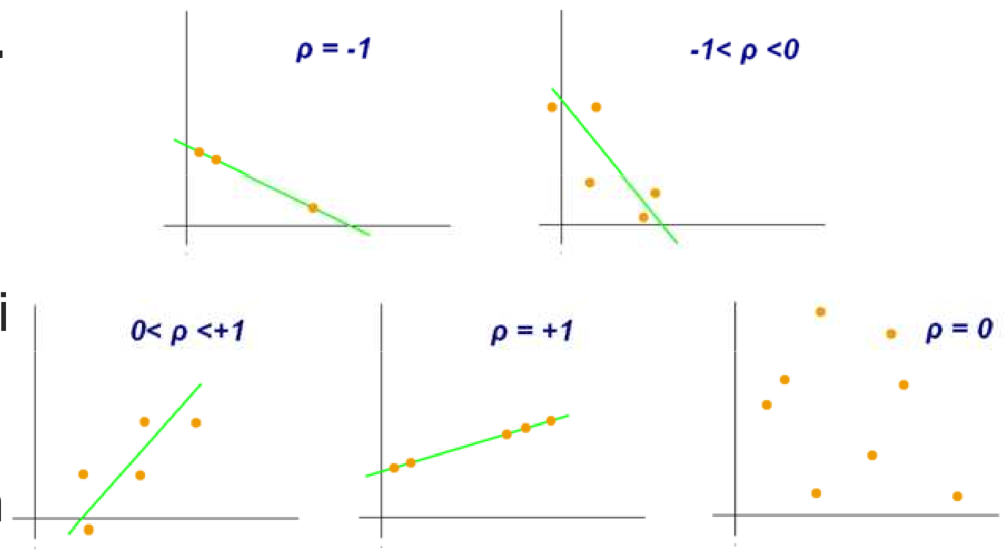
# PEARSONŮV KORELAČNÍ KOEFICIENT

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

$$r = \frac{\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})(y_n - \bar{y})}{\sigma_x \, \sigma_y} = \frac{1}{N}\sum_{n=1}^{N}\frac{(x_n - \bar{x})}{\sigma_x} \cdot \frac{(y_n - \bar{y})}{\sigma_y}$$

nabývá hodnoty v intervalu ⟨-1, 1⟩.

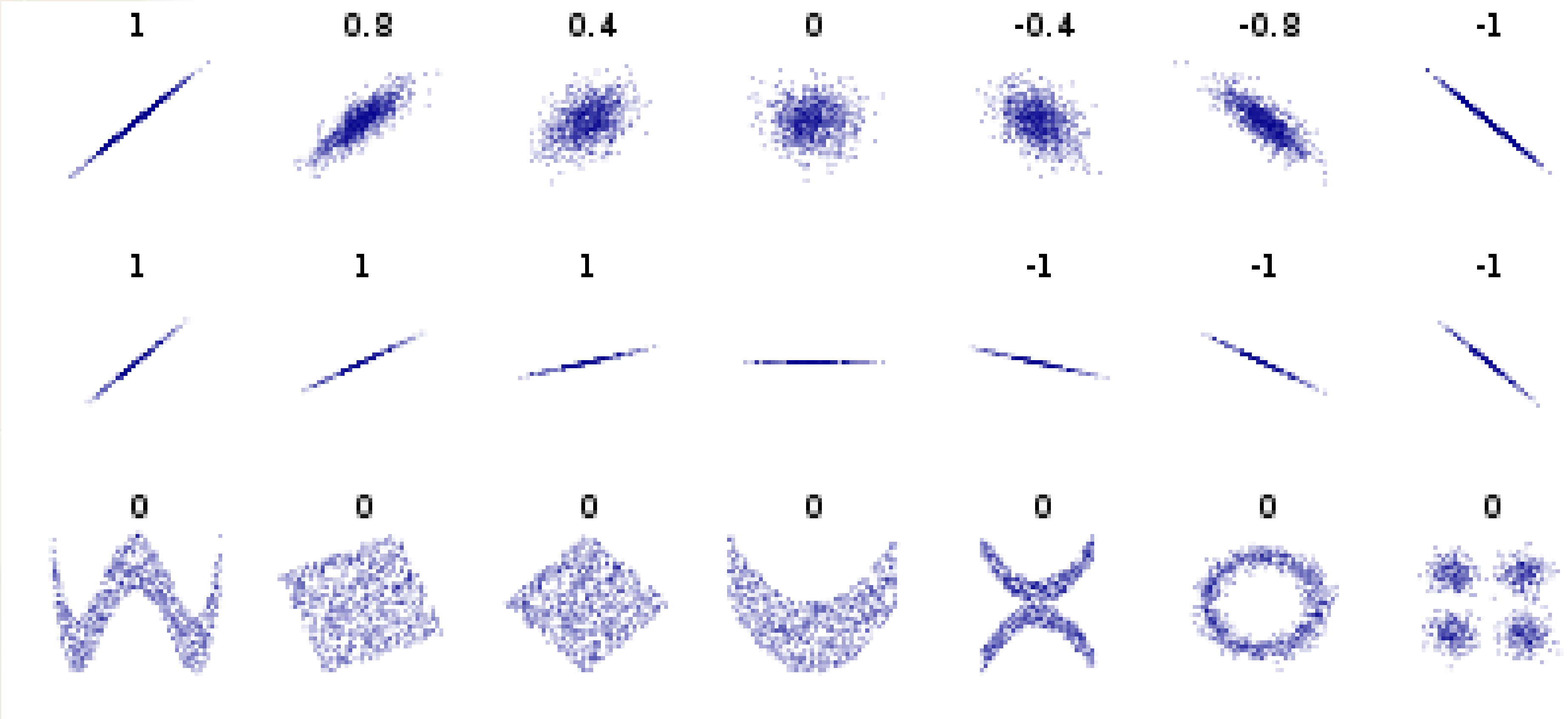
Když je roven +1, znamená to zcela pozitivní (přímou) **lineární** korelaci (y=kx, tj. funkční závislost), při nulové hodnotě není mezi proměnnými **lineární** vztah a hodnota -1 znamená zcela negativní **lineární** korelaci mezi oběma proměnnými (y=-kx) .




By Kiatdd - Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=37108966

# PEARSONŮV KORELAČNÍ KOEFICIENT

$$r = \frac{\frac{1}{N}\sum_{n=1}^{N}(x_n-\bar{x})(y_n-\bar{y})}{\sigma_x\,\sigma_y} = \frac{1}{N}\sum_{n=1}^{N}\frac{(x_n-\bar{x})}{\sigma_x}\cdot\frac{(y_n-\bar{y})}{\sigma_y}$$



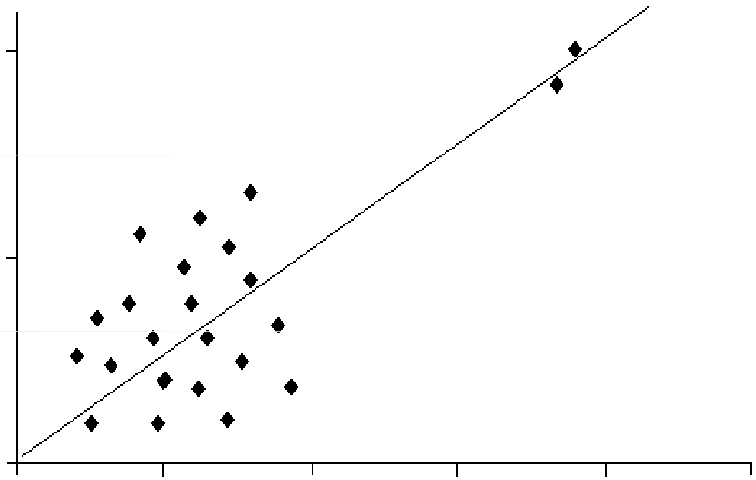
# PEARSONŮV KORELAČNÍ KOEFICIENT

- nabývá hodnot od −1 do +1, které značí perfektní lineární vztah (záporný nebo kladný)
  - v případě kladné korelace hodnoty obou proměnných zároveň stoupají;
  - v případě záporné korelace hodnota jedné proměnné stoupá a druhé klesá;
  - v případě neexistence lineárního vztahu $r = 0$;
- je nezávislý na jednotkách původních proměnných, je bezrozměrný;
- při změně pořadí proměnných se výše korelačního koeficientu nemění;
- korelační koeficient je platný pouze v rozmezí daném použitými daty;
- korelační koeficient výrazně odlišný od nuly není důkazem funkčního vztahu proměnných, jiného než lineárního;
- malá hodnota korelačního koeficientu není známkou nefunkčního vztahu proměnných.

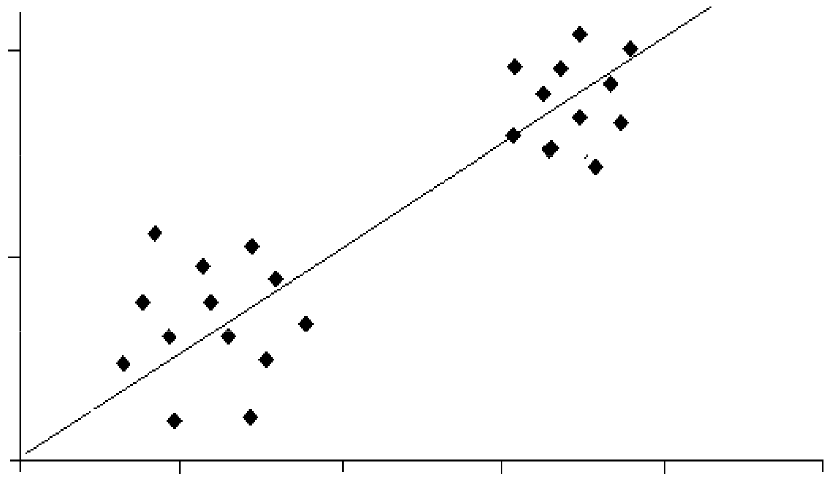
# PEARSONŮV KORELAČNÍ KOEFICIENT

Správná interpretace Pearsonova korelačního koeficientu předpokládá, že obě proměnné jsou náhodné veličiny a mají společné dvourozměrné normální rozdělení. Potom nulový korelační koeficient znamená, že veličiny jsou i statisticky **nezávislé**.

Pokud není splněn předpoklad dvourozměrné normality, z nulové hodnoty korelačního koeficientu nelze usuzovat na nic víc, než že veličiny jsou nekorelované.



odlehlé hodnoty

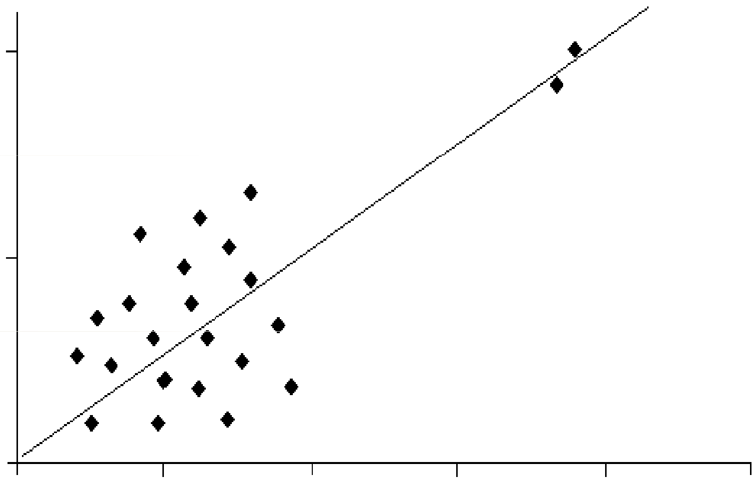


bimodální rozložení

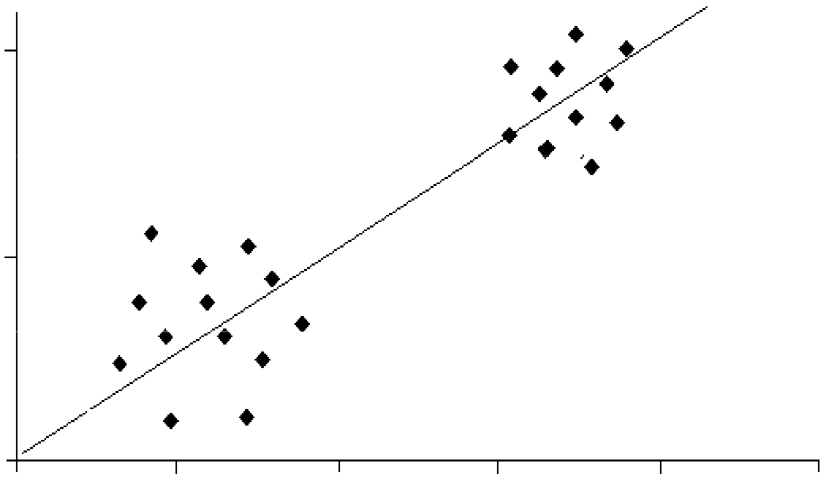
# PEARSONŮV KORELAČNÍ KOEFICIENT

může být nadhodnocen:

- ☑ vlivem třetí skryté proměnné
- ☑ přítomností odlehlých hodnot
- ☑ data jsou složena z různých podskupin (tříd)



odlehlé hodnoty



bimodální rozložení

# SPEARMANŮV KORELAČNÍ KOEFICIENT

$$r_s = \rho_{\mathrm{rg}_X, \mathrm{rg}_Y} = \frac{\mathrm{cov}(\mathrm{rg}_X, \mathrm{rg}_Y)}{\sigma_{\mathrm{rg}_X} \sigma_{\mathrm{rg}_Y}}$$

Hodnotí, jak dobře lze vztah mezi dvěma proměnnými popsat pomocí **monotónní** funkce.

Pokud jsou všechna pořadí určena ***různými celými čísly od 1 do n***, pak jeho hodnotu lze určit pomocí vztahu

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}.$$

kde $d_i = rg(X_i) - rg(Y_i)$ je rozdíl mezi pořadími každého pozorování

# SPEARMANŮV KORELAČNÍ KOEFICIENT

Určeme korelaci mezi IQ sledovaných osob a počtem hodin, které stráví za týden sledováním televize.

(https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient)

| $X_i$ IQ | $Y_i$ hours |
|---|---|
| 86 | 0 |
| 112 | 6 |
| 106 | 7 |
| 113 | 12 |
| 110 | 17 |
| 97 | 20 |
| 100 | 27 |
| 99 | 28 |
| 103 | 29 |
| 101 | 50 |

| $X_i$ IQ | $Y_i$ hours | rank $x_i$ | rank $y_i$ | $d_i$ | $d_i^2$ |
|---|---|---|---|---|---|
| 86 | 0 | 1 | 1 | 0 | 0 |
| 97 | 20 | 2 | 6 | −4 | 16 |
| 99 | 28 | 3 | 8 | −5 | 25 |
| 100 | 27 | 4 | 7 | −3 | 9 |
| 101 | 50 | 5 | 10 | −5 | 25 |
| 103 | 29 | 6 | 9 | −3 | 9 |
| 106 | 7 | 7 | 3 | 4 | 16 |
| 110 | 17 | 8 | 5 | 3 | 9 |
| 112 | 6 | 9 | 2 | 7 | 49 |
| 113 | 12 | 10 | 4 | 6 | 36 |

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}.$$

$$r_s = 1 - \frac{6 \times 194}{10.(10^2 - 1)} = -0.1757575...$$

# SPEARMANŮV KORELAČNÍ KOEFICIENT

## ZÁVĚR

Nízká hodnota korelačního koeficientu ukazuje, že korelace mezi IQ a hodinami sledování TV je malá a jeho záporná hodnota indikuje, že delší čas před TV je vázán na menší hodnotu IQ.

Nehodnotíme zde ovšem statistickou signifikanci neboli riziko falešně pozitivního výsledku korelačního testu…

# SPEARMANŮV KORELAČNÍ KOEFICIENT
## PŘÍKLADY

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}.$$

| $X_i$ | $Y_i$ | rank $x_i$ | rank $y_i$ | $d_i$ | $d_i^2$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 |
| 2 | 2 | 2 | 2 | 0 | 0 |
| 3 | 3 | 3 | 3 | 0 | 0 |
| 4 | 4 | 4 | 4 | 0 | 0 |

$$r_s = 1 - \frac{6 \times 0}{4 \times (16-1)} = 1$$

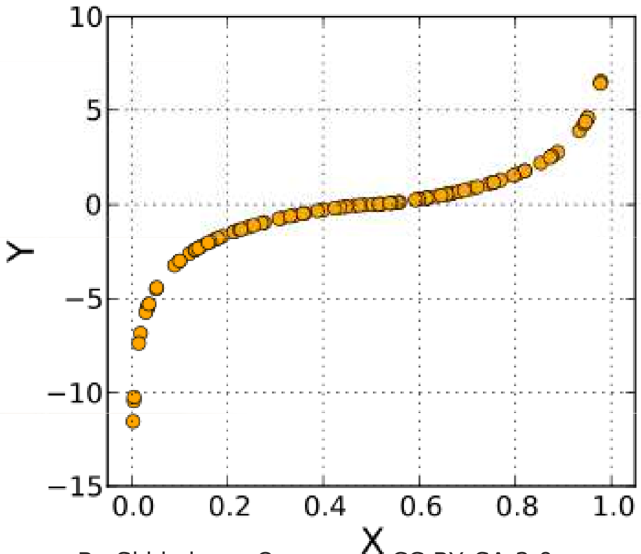
| $X_i$ | $Y_i$ | rank $x_i$ | rank $y_i$ | $d_i$ | $d_i^2$ |
|---|---|---|---|---|---|
| 1 | -1 | 1 | 4 | -3 | 9 |
| 2 | -2 | 2 | 3 | -1 | 1 |
| 3 | -3 | 3 | 2 | 1 | 1 |
| 4 | -4 | 4 | 1 | 3 | 9 |

$$r_s = 1 - \frac{6 \times 20}{4 \times (16-1)} =$$

$$= 1 - \frac{120}{60} = -1$$

# SPEARMANŮV KORELAČNÍ KOEFICIENT

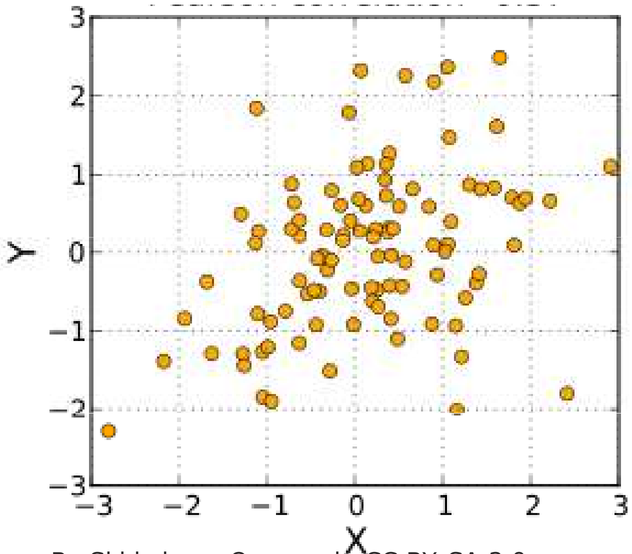
Spearmanův koeficient=1
Pearsonův koeficient=0,88




By Skbkekas - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=8778554

Hodnota Spearmanova korelačního koeficientu je rovna 1 pokud jsou srovnávané veličiny monotónně závislé, i pokud je jejich vzájemný vztah nelineární.

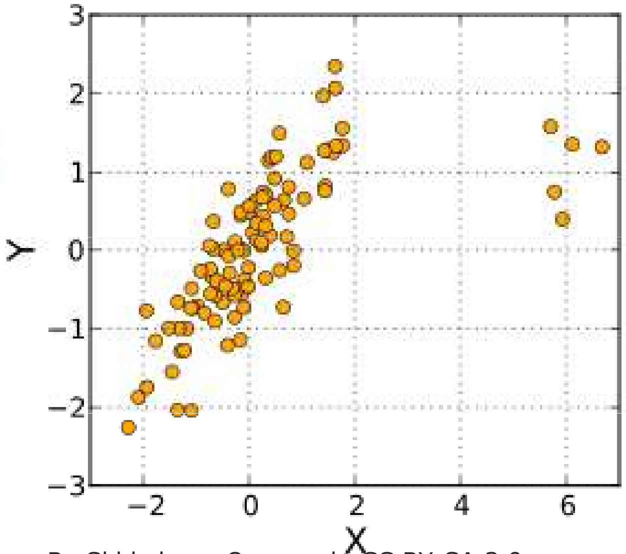
Spearmanův koeficient=0,35
Pearsonův koeficient=0,37




By Skbkekas - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=8778562

Pokud jsou data rozložena přibližně elipticky a nevyskytují se žádné významné odlehlé hodnoty, pak Pearsonův a Spearmanův korelační koeficient nabývají přibližně týchž hodnot.

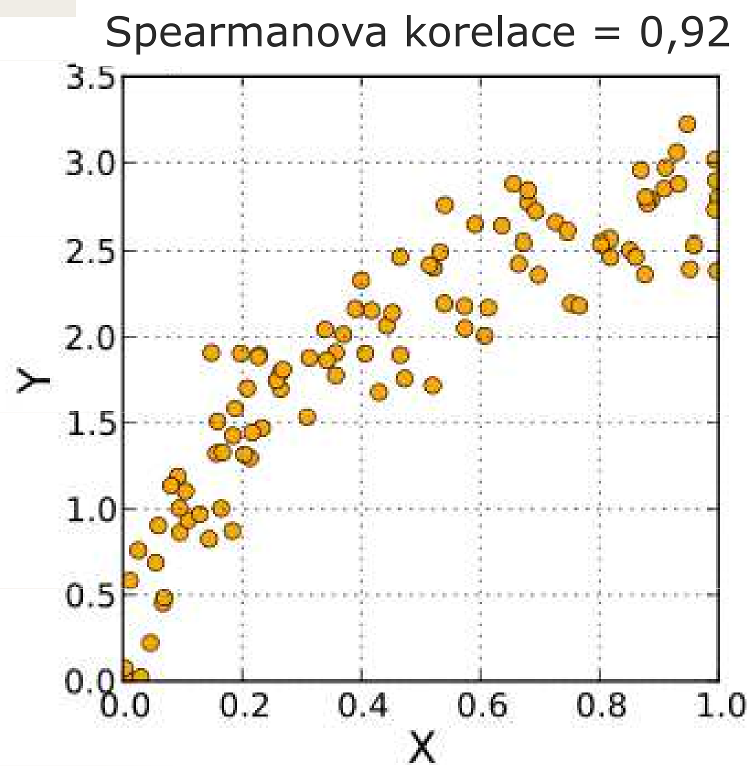
Spearmanův koeficient=0,84
Pearsonův koeficient=0,67



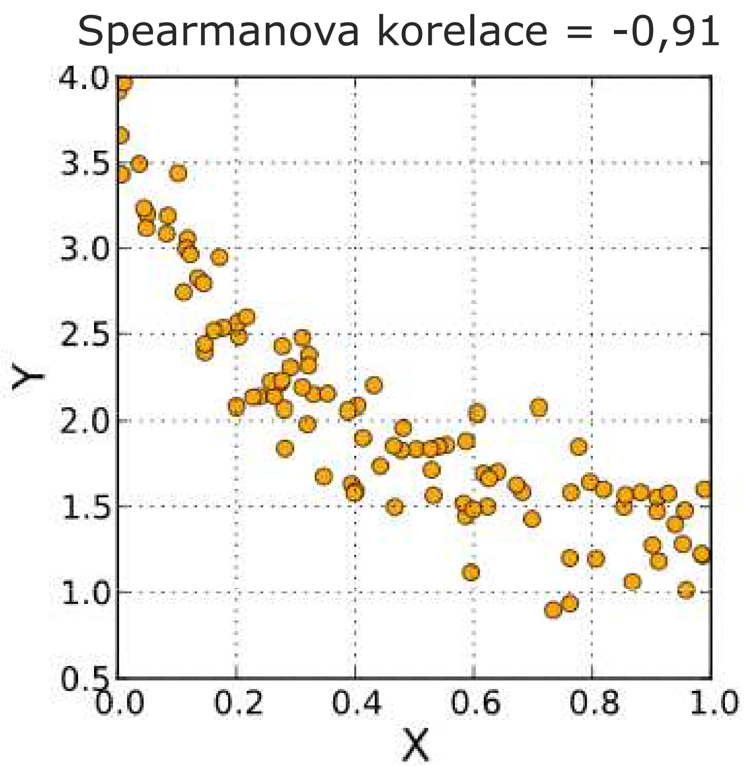

By Skbkekas - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=8778570

Spearmanův korelační koeficient je méně citlivý vůči významným odlehlým hodnotám. To je způsobeno skutečností, že Spearmanův koeficient redukuje hodnoty odlehlých hodnot na jejich pořadí v posloupnosti.

# SPEARMANŮV KORELAČNÍ KOEFICIENT








Kladná hodnota Spearmanova korelačního koeficientu odpovídá monotónně rostoucí závislosti mezi náhodnými veličinami X a Y.

Záporná hodnota Spearmanova koeficientu odpovídá monotónně klesající závislosti mezi X a Y.

# ČASOVĚ ZÁVISLÁ DATA

## KŘÍŽOVÁ (VZÁJEMNÁ) KORELAČNÍ POSLOUPNOST

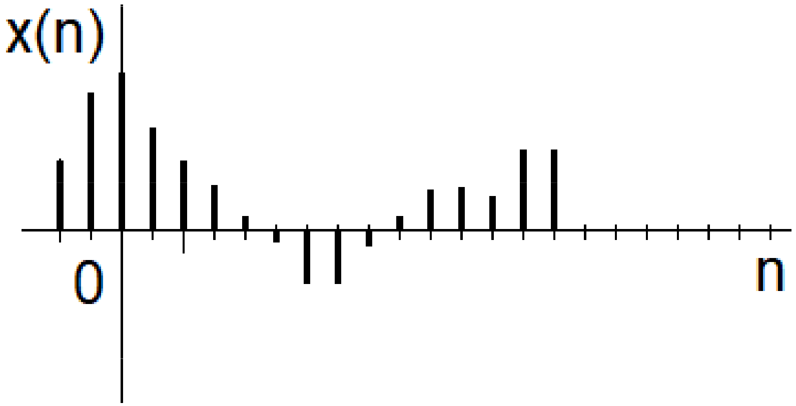
1. $$R_{x1x2}(mT_{vz}) = \sum_{-\infty}^{\infty} x_1(nT_{vz})x_2(nT_{vz} + mT_{vz})$$

2. $$R_{x1x2}(mT_{vz}) = \sum_{n}^{N} x_1(nT_{vz})x_2(nT_{vz} + mT_{vz})$$

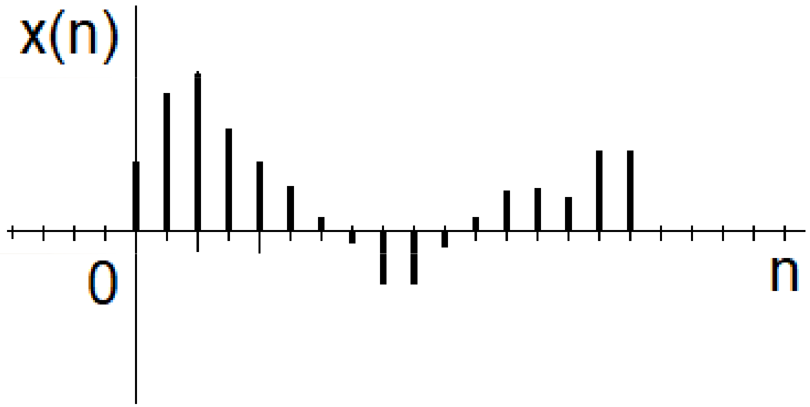
3. $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n + m) \qquad R_{x1x2}(m) = \sum^{N} x_1(n)x_2(n - m)$$

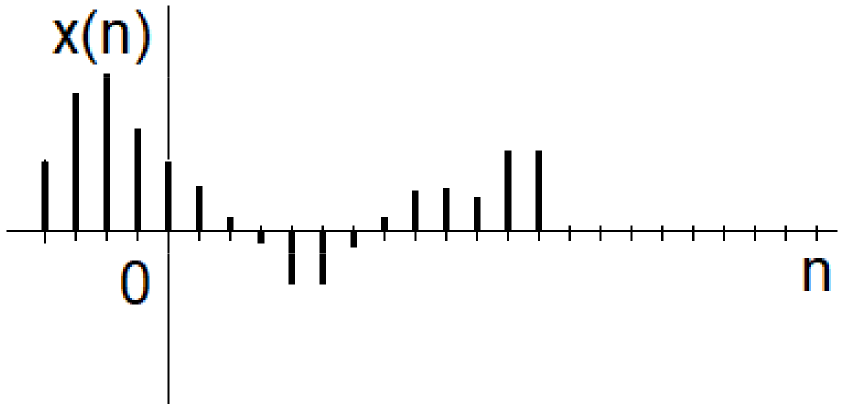
4. $$R_{x1x2}(m) = \frac{1}{N}\sum^{N} x_1(n)x_2(n + m) \qquad R_{x1x2}(m) = \frac{1}{N}\sum^{N} x_1(n)x_2(n - m)$$

# KŘÍŽOVÁ KORELAČNÍ POSLOUPNOST

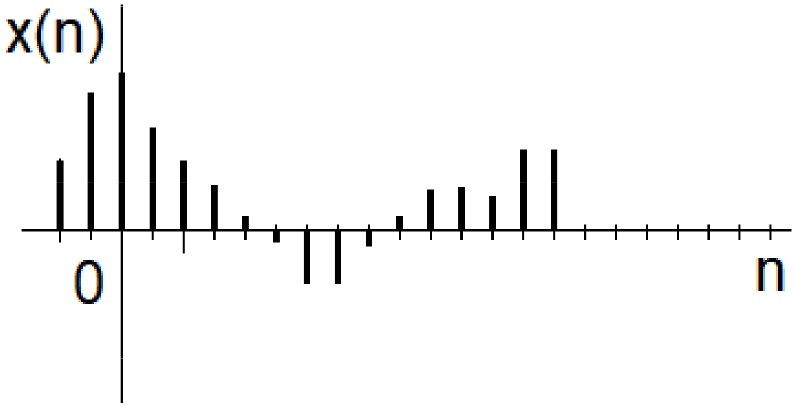



**? x(n+2) ?**

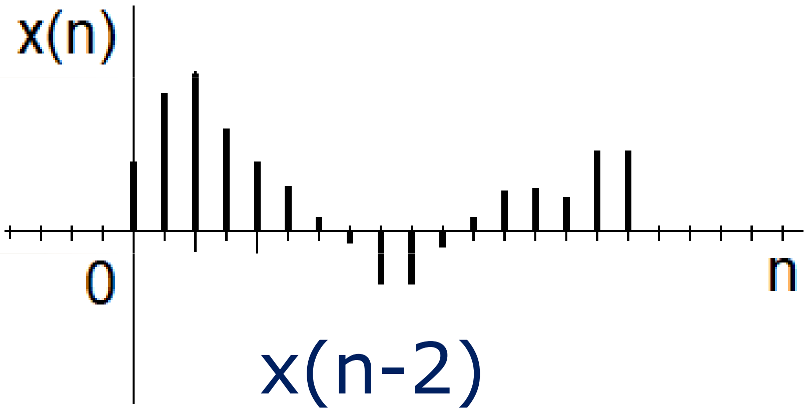


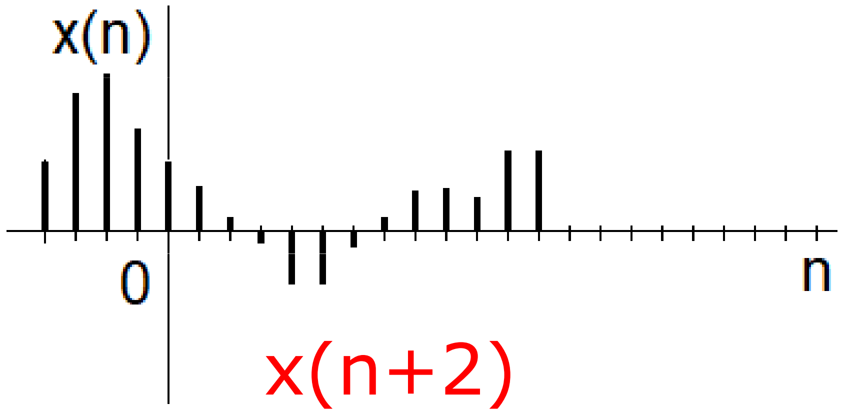



# KŘÍŽOVÁ KORELAČNÍ POSLOUPNOST




**? x(n+2) ?**



# ČASOVĚ ZÁVISLÁ DATA

## KŘÍŽOVÁ (VZÁJEMNÁ) KORELAČNÍ POSLOUPNOST

1. $$R_{x1x2}(mT_{vz}) = \sum_{-\infty}^{\infty} x_1(nT_{vz})x_2(nT_{vz} + mT_{vz})$$

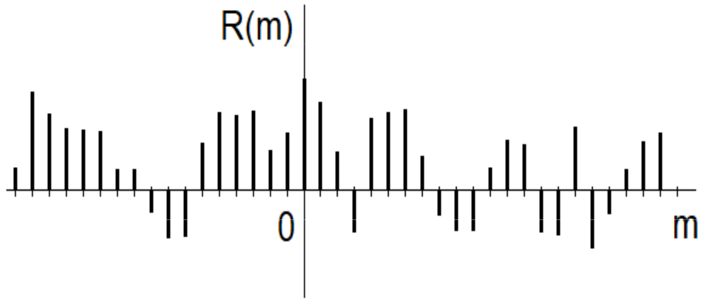
2. $$R_{x1x2}(mT_{vz}) = \sum_{n}^{N} x_1(nT_{vz})x_2(nT_{vz} + mT_{vz})$$

3. $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n + m) \qquad R_{x1x2}(m) = \sum^{N} x_1(n)x_2(n - m)$$

4. $$R_{x1x2}(m) = \frac{1}{N}\sum^{N} x_1(n)x_2(n + m) \qquad R_{x1x2}(m) = \frac{1}{N}\sum^{N} x_1(n)x_2(n - m)$$

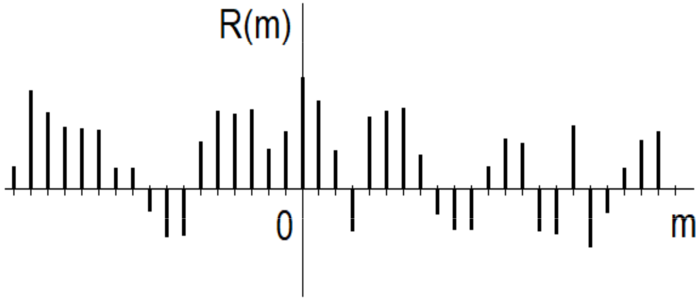
# KŘÍŽOVÁ KORELAČNÍ POSLOUPNOST

$$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n+m)$$ **?** $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n-m)$$

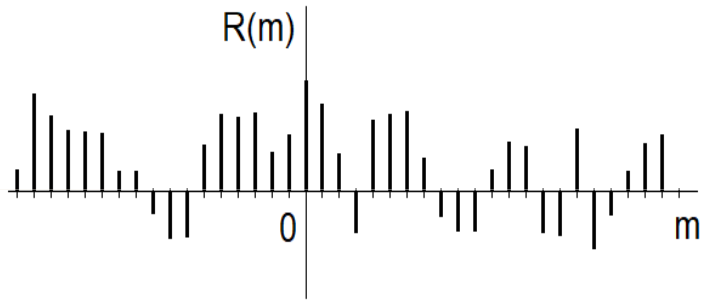


# KŘÍŽOVÁ KORELAČNÍ POSLOUPNOST

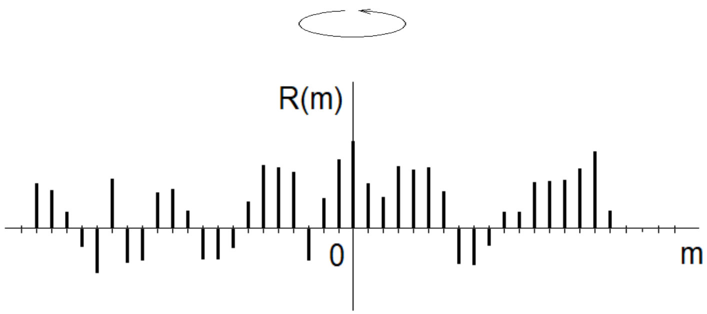
$$R_{x1x2}(m)=\sum_{n}^{N}x_1(n)x_2(n+m) \quad \textbf{?} \quad R_{x1x2}(m)=\sum_{n}^{N}x_1(n)x_2(n-m)$$



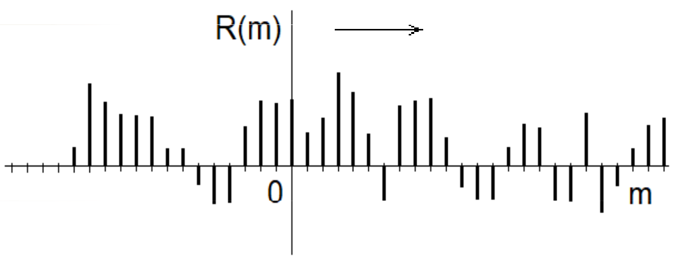

**1.**



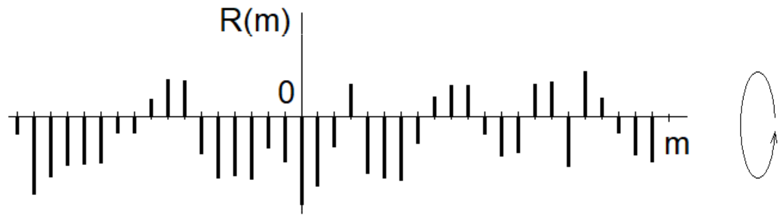

**2.**




**3.**

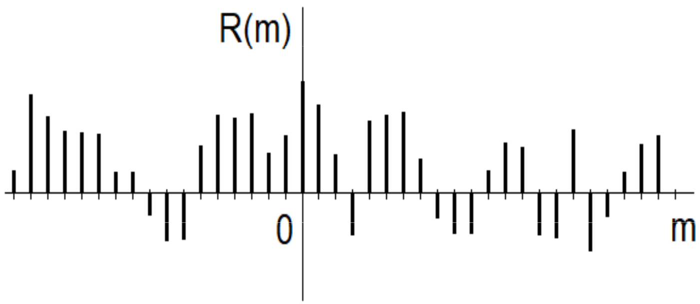



**4.**

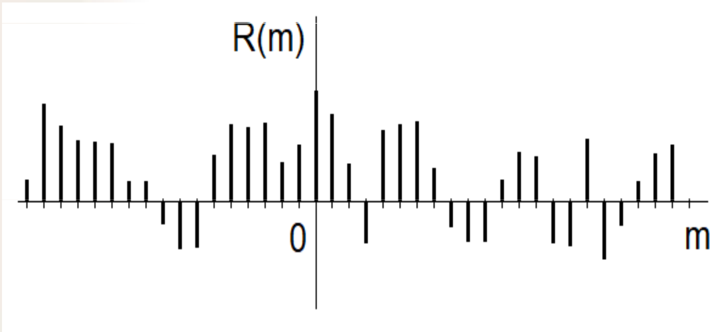


# KŘÍŽOVÁ KORELAČNÍ POSLOUPNOST

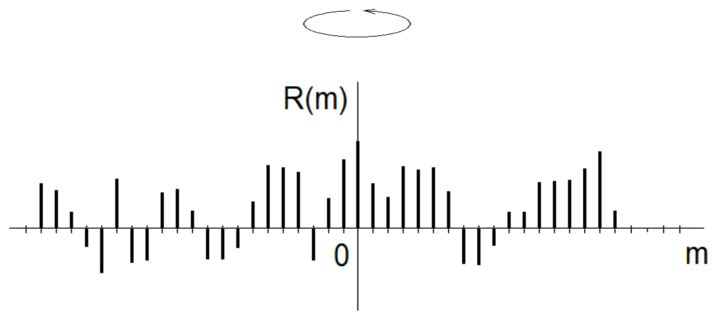
$$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n+m) \quad \textbf{?} \quad R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n-m)$$



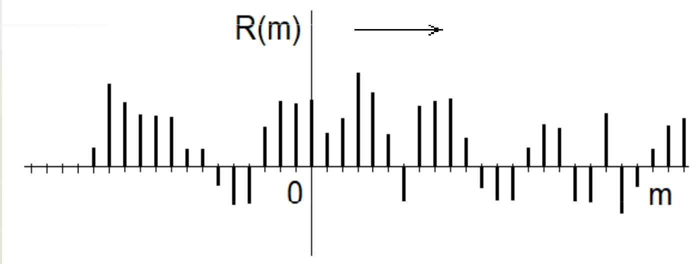

1.



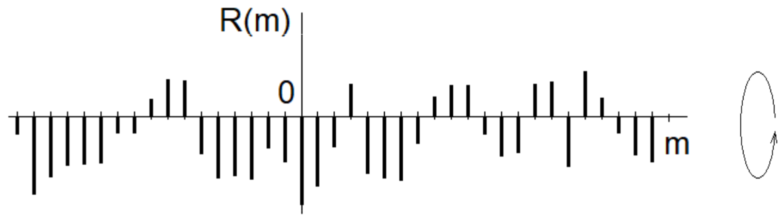

2.




3.




4.



# KORELAČNÍ POSLOUPNOSTI

$$R_{x1x2}(m) = \sum_{n}^{N} x_1(n) x_2(n+m)$$

## KŘÍŽOVÁ (VZÁJEMNÁ) KORELAČNÍ POSLOUPNOST

$$x_1(n) \neq x_2(n)$$

## AUTOKORELAČNÍ POSLOUPNOST

$$x_1(n) = x_2(n)$$

## KŘÍŽOVÁ KOVARIANČNÍ/AUTOKOVARIANČNÍ POSLOUPNOST

$$x_1(n) = x_{1orig}(n) - m_{x1orig} \qquad x_2(n) = x_{2orig}(n) - m_{x2orig}$$

# KORELAČNÍ POSLOUPNOSTI

## VLASTNOSTI VZÁJEMNÉ KORELAČNÍ POSLOUPNOSTI

- ☑ vzájemná korelační posloupnost není komutativní, tj.

$$R_{x1x2}(n) \neq R_{x2x1}(n)$$

## VLASTNOSTI AUTOKORELAČNÍ POSLOUPNOSTI

- ☑ sudá, tj. $R(-n) = R(n)$;
- ☑ $\forall \tau \in \mathbb{Z}$: $R(0) \geq R(n)$;
- ☑ $R(0)$ je rovna energii, resp. výkonu posloupnosti $x(n)$;
- ☑ pokud je posloupnost $x(n)$ periodická, pak je její autokorelační posloupnost rovněž periodická s toutéž periodou.

## VLASTNOSTI AUTOKOVARIANČNÍ POSLOUPNOSTI

- ☑ $R(0)$ je rovna disperzi posloupnosti $x(n)$ pokud je určena pomocí vztahu

$$R_{xx}(m) = \frac{1}{N}\sum_{n}^{N}[x(n)-m_x][x(n+m)-m_x]$$

# KORELAČNÍ POSLOUPNOSTI

$$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n+m) \quad \textbf{?} \quad R_{x1x2}(m) = \frac{1}{N}\sum_{n}^{N} x_1(n)x_2(n+m)$$

# ENERGIE & VÝKON

$$E_d = \sum_{n}^{N} x^2(nT_{vz})$$

$$P_{dn} = \frac{1}{N}\sum_{n}^{N} x^2(n)$$

# KORELAČNÍ POSLOUPNOST VS. PERSONŮV KORELAČNÍ KOEFICIENT

## ☑ VZÁJEMNÁ KORELAČNÍ POSLOUPNOST

$$\hat{R}_{x1x2}(m) = \frac{1}{N}\sum_{n=0}^{N-1} x_1(n)x_2(n+m)$$

$$\hat{R}_{x1x2}(0) = \frac{1}{N}\sum_{n=0}^{N-1} x_1(n)x_2(n+0) = \frac{1}{N}\sum_{n=0}^{N-1} x_1(n)x_2(n)$$

## ☑ PEARSONŮV KORELAČNÍ KOEFICIENT

$$r = \frac{\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})(y_n - \bar{y})}{\sigma_x \, \sigma_y} = \frac{1}{N}\sum_{n=1}^{N} \frac{(x_n - \bar{x})}{\sigma_x} \cdot \frac{(y_n - \bar{y})}{\sigma_y}$$

Pokud jsou srovnávané posloupnosti $x_1(n)$ a $x_2(n)$ standardizovány, pak každý vzorek jejich vzájemné korelační posloupnosti je roven odpovídající hodnotě Pearsonova korelačního koeficientu.

# KORELAČNÍ POSLOUPNOST VS. KONVOLUCE

vzájemná korelační posloupnost:

$$R_{x1x2}(m) = \sum_{n}^{N} x_1(n) x_2(n-m)$$

konvoluce:

$$y(n) = x_1(n) * x_2(n) = \sum_{m} x_1(m).x_2(n-m) = \sum_{m} x_1(n-m).x_2(m)$$

Po záměně symbolů

$$y(m) = x_1(m) * x_2(m) = \sum_{n} x_1(n).x_2(m-n)$$

Vztahy by byly ekvivalentní, pokud by posloupnost $x_2$ byla invertována v čase, tj. když by platilo $x_2(n-m)$.

# PARCIÁLNÍ KORELACE

**Parciální korelační koeficient** je míra vztahu mezi dvěma veličinami, které jsou řízeny jednou nebo více řídicími veličinami.

# PARCIÁLNÍ KORELACE

**PŘÍKLAD:** (https://en.wikipedia.org/wiki/Partial_correlation)

Předpokládejme, že známe hodnoty tří proměnných, *X*, *Y*, a *Z*:

| X | Y | Z |
|---|---|---|
| 2 | 1 | 0 |
| 4 | 2 | 0 |
| 15 | 3 | 1 |
| 20 | 4 | 1 |

Pokud *Z* = 0, jsou hodnoty *X* rovny přesně dvounásobku *Y*, a pokud *Z* = 1, *X* je přesně rovno pětinásobku *Y*. Tak v závislosti na Y je přesný vztah mezi *X* a *Y*; tato relace ale nemůže být přesná bez reference k hodnotám *Z*.

Pokud určíme Pearsonův korelační koeficient mezi posloupnostmi *X* a *Y*, je výsledná hodnota 0,836. Určíme-li ale parciální korelaci mezi *X* a *Y* pomocí dále uvedeného vztahu, odpovídá její hodnota 0,919 podstatně silnějšímu vzájemnému vztahu.

# PARCIÁLNÍ KORELACE

## JAK JI SPOČÍTAT?

Parciální korelace $r_{XY.\mathbf{Z}}$ mezi posloupnostmi X a Y za předpokladu, že existují řídicí veličiny $\mathbf{Z} = \{Z_1, Z_2, \ldots, Z_n\}$ je rovna Pearsonově korelaci mezi rezidui $e_x$ a $e_y$, která plynou z lineární regrese posloupnosti X a proměnných **Z**, resp. Y a **Z**.

- ☑ pomocí lineární regrese;
- ☑ pomocí maticové inverze;
- ☑ pomocí rekurzivního vztahu

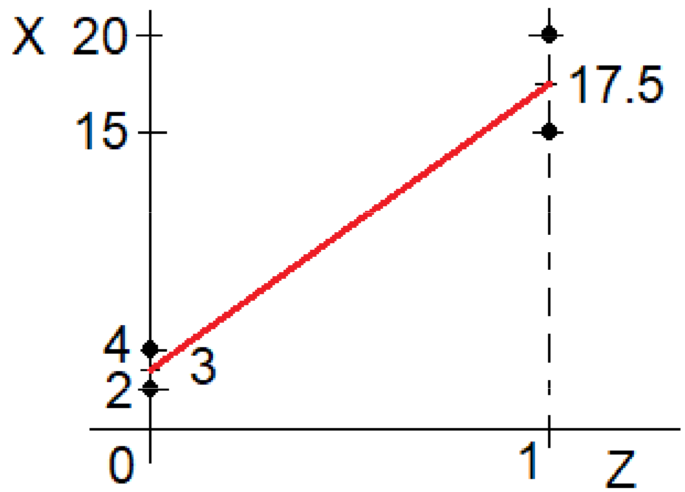
s jedinou řídicí proměnnou se tento vztah redukuje na tvar

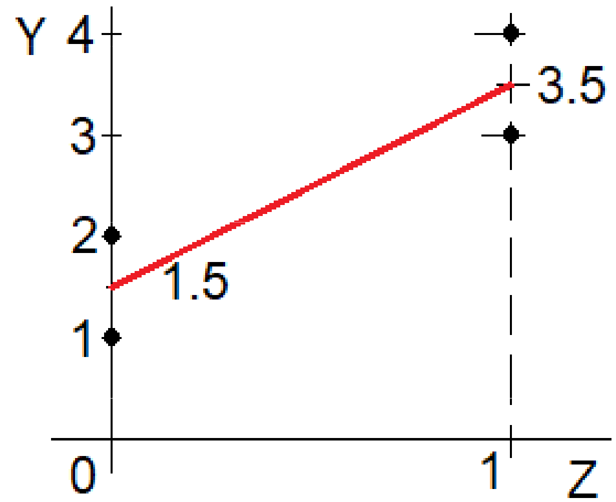
$$r_{XYZ} = \frac{r_{XY} - r_{XZ} r_{YZ}}{\sqrt{(1 - r_{XZ}{}^2)(1 - r_{YZ}{}^2)}}$$

# PARCIÁLNÍ KORELACE

## JAK JI SPOČÍTAT?

| X | Y | Z |
|---|---|---|
| 2 | 1 | 0 |
| 4 | 2 | 0 |
| 15 | 3 | 1 |
| 20 | 4 | 1 |




| X | ΔX | Y | ΔY |
|---|---|---|---|
| 2 | -1.0 | 1 | -0.5 |
| 4 | 1.0 | 2 | 0.5 |
| 15 | -2.5 | 3 | -0.5 |
| 20 | 2.5 | 4 | 0.5 |

$$r_{\Delta X \Delta Y} = 0{,}919$$

# PARCIÁLNÍ AUTOKORELAČNÍ POSLOUPNOST

**Parciální autokorelační posloupnost** (**PACF**) určuje hodnoty parciální korelace mezi hodnotami časové řady a jejími zpožděnými hodnotami, přičemž jako řídící posloupnost je použita posloupnost pro všechna menší zpoždění.

Na rozdíl od obyčejné autokorelační posloupnosti, která již na ničem dalším nezávisí.

Je-li známa časová řada x(n), pak parciální autokorelační posloupnost $R_{xx}(k)$ pro zpoždění k je určena podle následujících vztahů

$$R_{xx}(1) = \mathrm{Cor}\,(x(n), x(n+1)),$$

$$R_{xx}(k) = \mathrm{Cor}\,(x(n) - \hat{x}(n), x(n+k) - \hat{x}(n+k)), k = 2,3, \ldots,$$

kde $\hat{x}(n)$ a $\hat{x}(n+k)$ jsou hodnoty nejlepší lineární predikce x(n) a x(n+k) z hodnot x(n+1), …, x(n+k-1).

# ODHAD KORELAČNÍ POSLOUPNOSTI

## KORELAČNÍ POSLOUPNOST

**1.** $$R_{x1x2}(mT_s) = \sum_{n}^{N} x_1(nT_s)x_2(nT_s + mT_s)$$

**2.** $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n+m)$$ $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n-m)$$

**3.** $$R_{x1x2}(m) = \frac{1}{N}\sum_{n}^{N} x_1(n)x_2(n+m)$$ $$R_{x1x2}(m) = \frac{1}{N}\sum_{n}^{N} x_1(n)x_2(n-m)$$

# ODHAD KORELAČNÍ POSLOUPNOSTI

## KORELAČNÍ POSLOUPNOST

**1.** $$R_{x1x2}(mT_s) = \sum_{n}^{N} x_1(nT_s)x_2(nT_s + mT_s)$$

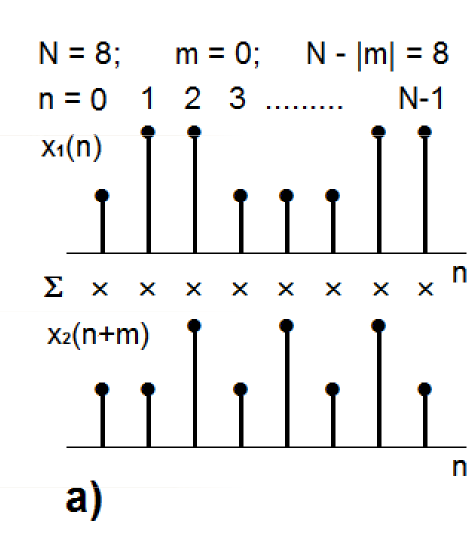
**2.** $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n+m)$$ $$R_{x1x2}(m) = \sum_{n}^{N} x_1(n)x_2(n-m)$$

**3.** $$\boxed{R_{x1x2}(m) = \frac{1}{N}\sum_{n}^{N} x_1(n)x_2(n+m) \qquad R_{x1x2}(m) = \frac{1}{N}\sum_{n}^{N} x_1(n)x_2(n-m)}$$

# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o téže délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

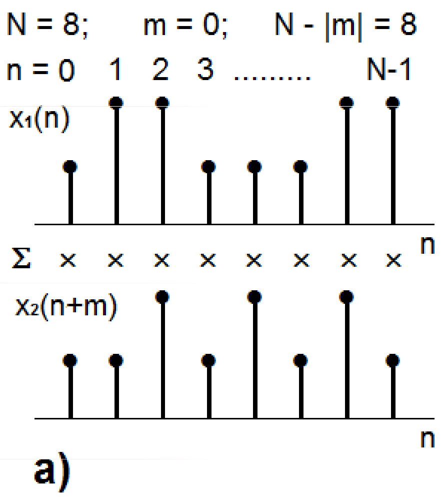
$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$

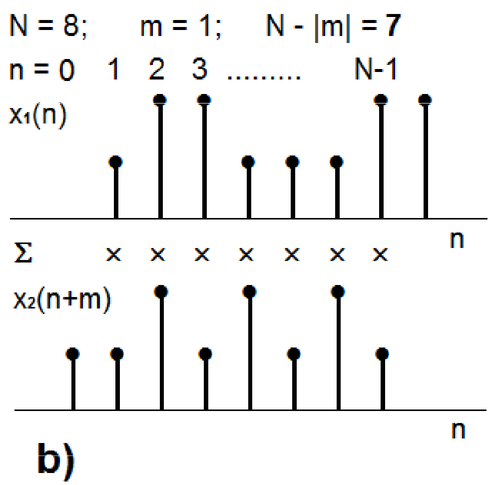


# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o týž délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$



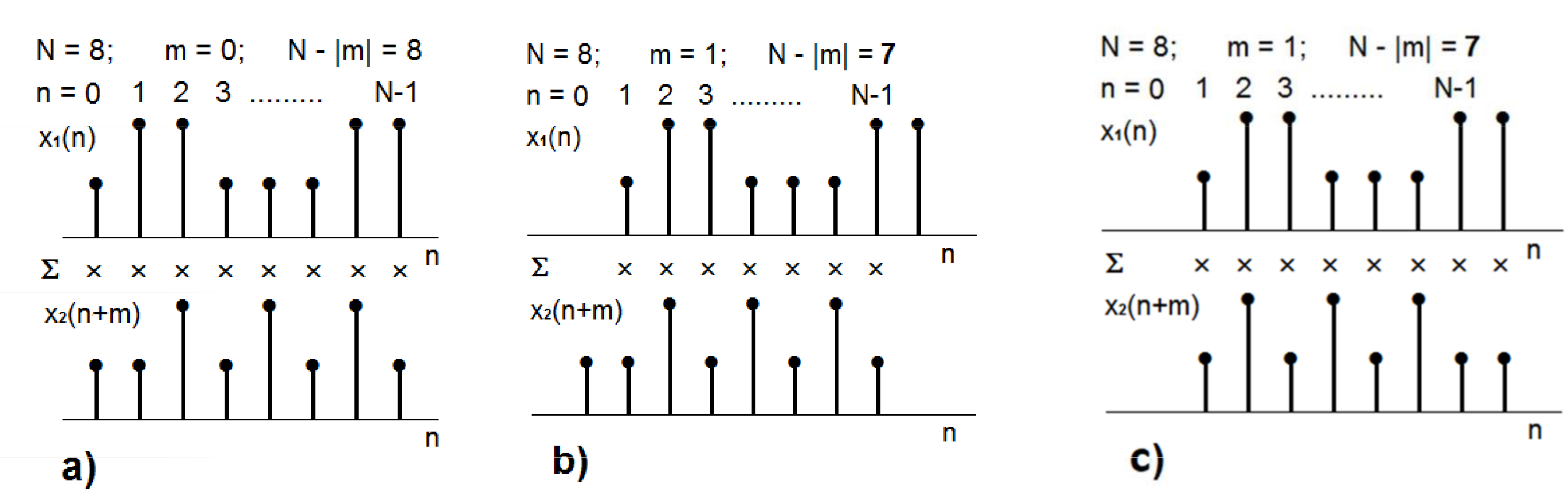




?

# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o týž délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$

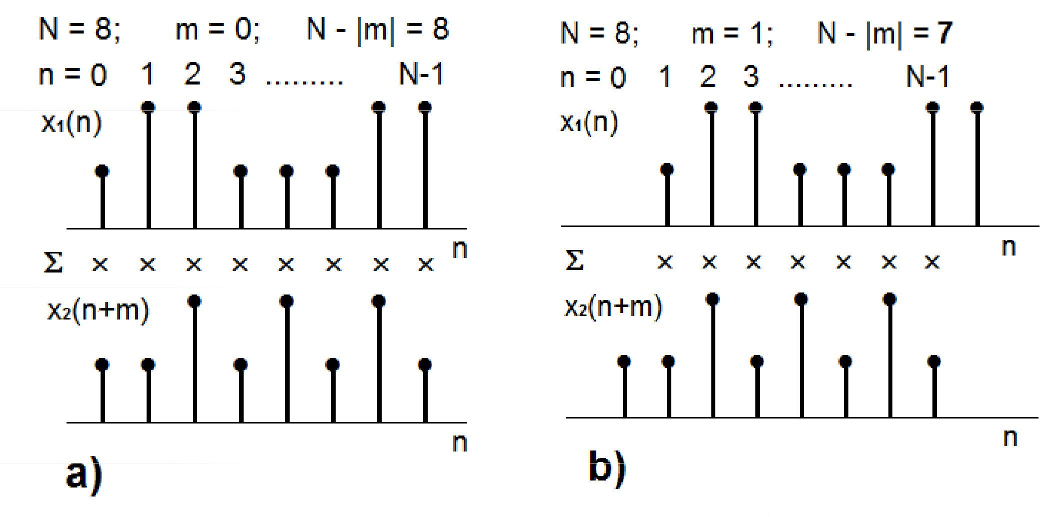



**kruhová korelace**

# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o týž délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$

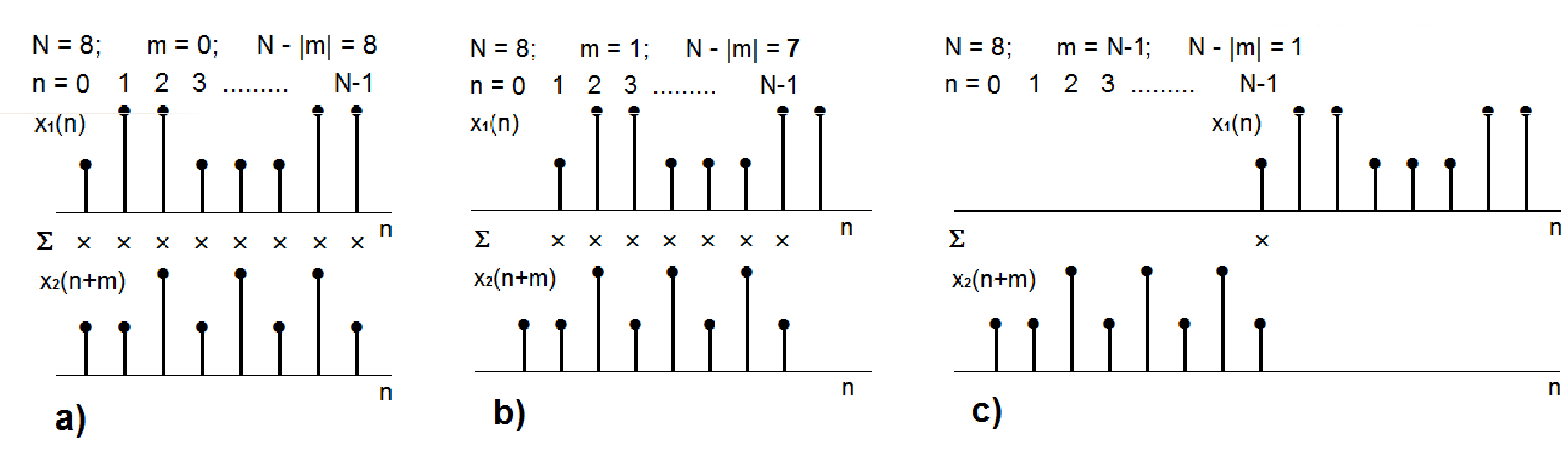



$$A = \frac{1}{N} \text{ nebo } A = \frac{1}{N-1}$$

# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o týž délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$

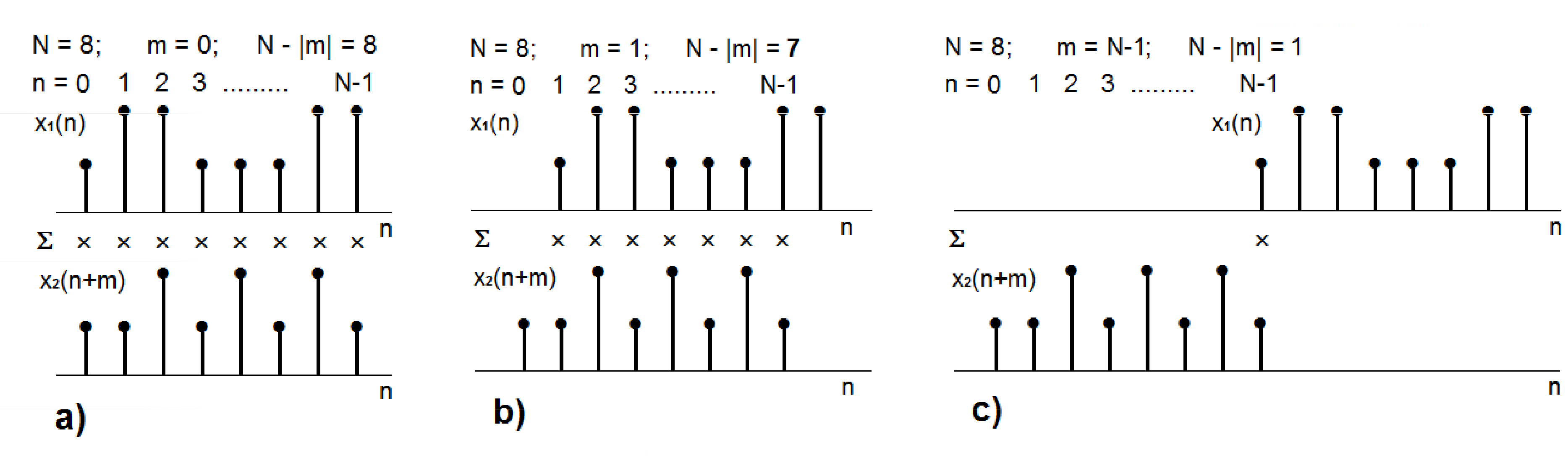



$$A = \frac{1}{N} \text{ nebo } A = \frac{1}{N-1}$$

# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o týž délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$




$$_1\tilde{r}_{x1x2}(m) = \frac{1}{N-|m|}\sum_{n=0}^{N-|m|-1} x_1(n)x_2(n+m); \quad m = 0,1,\ldots,N-1 \qquad _2\tilde{r}_{x1x2}(m) = \frac{1}{N}\sum_{n=0}^{N-|m|-1} x_1(n)x_2(n+m); \quad m = 0,1,\ldots,N-1$$

# KRÁTKÝ ODHADOVÝ EXKURZ

- ☑ odhad parametru je závislý na volbě úseku analyzované veličiny;
- ☑ protože je výběr intervalu náhodný, je i odhad parametru náhodnou veličinou;
- ☑ základní (požadované) vlastnosti odhadů:
  - ➔ **nestrannost** – záruka, že v průměru se bude odhad pohybovat kolem správné hodnoty parametru

    $E\hat{q} = q; \quad \lim_{K \to \infty} E\hat{q} = q$ asymptoticky nestranný odhad

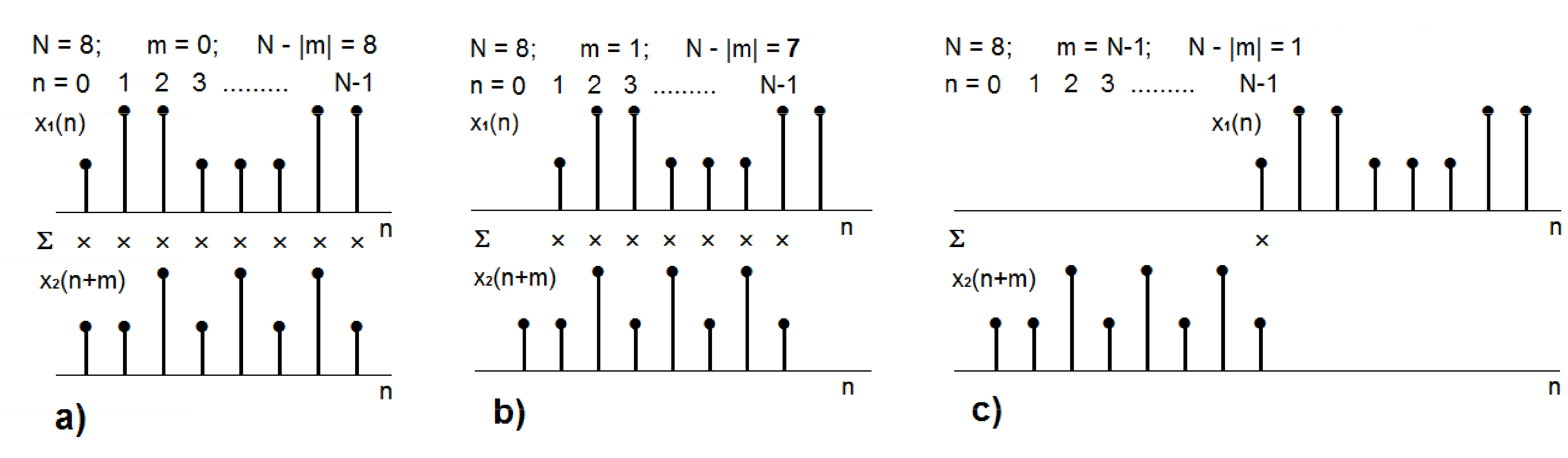
  - ➔ **konzistence** – čím delší bude zkoumaný interval, tím více se bude odhad blížit neznámé hodnotě

    $$(\forall \varepsilon > 0) \lim_{K \to \infty} P(\|\hat{q} - q\| > \varepsilon) = 0$$

# ODHAD KORELAČNÍ POSLOUPNOSTI

Nyní předpokládejme dvě konečné časové řady (obě o týž délce N vzorků), jejichž vzájemnou korelační posloupnost chceme spočítat.

$$\tilde{r}_{x1x2}(m) = A.\sum_{n=0}^{N-1} x_1(n).x_2(n+m) \qquad A = \frac{1}{N}$$




**1)**

$${}_1\tilde{r}_{x1x2}(m) = \frac{1}{N-|m|}\sum_{n=0}^{N-|m|-1} x_1(n)x_2(n+m); \quad m = 0,1,\ldots,N-1$$

**2)**

$${}_2\tilde{r}_{x1x2}(m) = \frac{1}{N}\sum_{n=0}^{N-|m|-1} x_1(n)x_2(n+m); \quad m = 0,1,\ldots,N-1$$

# ODHADY AUTOKORELAČNÍ POSLOUPNOSTI

ad1) **střední hodnota**

$$E\left[\tilde{r}_{xx1}(mT_{vz})\right] = \frac{1}{N-|m|} \cdot \sum_{n=0}^{N-|m|-1} \underbrace{E\left[x(nT_{vz})x(nT_{vz}+mT_{vz})\right]}_{} = \gamma_{xx}(mT_{vz})$$

takhle je definovaná AKF stacionárního diskrétního náhodného procesu

tzn. $\tilde{r}_{xx1}(mT_{vz})$ je nestranný odhad

**rozptyl** [Jenkins, G.M., Watts,D.G.: Spectral Analysis & Its Application, Holden-Day, 1968]

$$\operatorname{var}\left[\tilde{r}_{xx1}(mT_{vz})\right] \approx \frac{N}{(N-|m|)^2} \sum_{n=-\infty}^{\infty} \left[ |\gamma_{xx}(nT_{vz})|^2 + \gamma_{xx}(nT_{vz}-mT_{vz})\gamma_{xx}(nT_{vz}+mT_{vz}) \right]$$

$$\text{je-li} \quad \sum_{n=-\infty}^{\infty} |\gamma_{xx}(nT_{vz})|^2 < \infty, \text{ je} \quad \lim_{N\to\infty} \operatorname{var}\left[\tilde{r}_{xx1}(mT_{vz})\right] = 0$$

# ODHADY AUTOKORELAČNÍ POSLOUPNOSTI

ad1) pokračování

Protože $E[\tilde{r}_{xx1}(mT_{vz})] = \gamma_{xx}(mT_{vz})$ a $\lim_{N\to\infty} var[\tilde{r}_{xx1}(mT_{vz})] = 0$
je odhad $\tilde{r}_{xx1}(mT_{vz})$ konzistentní

pro velké hodnoty m má odhad $\tilde{r}_{xx1}(mT_{vz})$ velký rozptyl

# ODHADY AUTOKORELAČNÍ POSLOUPNOSTI

ad2) **střední hodnota**

$$E\left[\tilde{r}_{xx2}(mT_{vz})\right] = \frac{1}{N}\cdot\sum_{n=0}^{N-|m|-1} E\left[x(nT_{vz})x(nT_{vz}+mT_{vz})\right] =$$

$$= \frac{N-|m|}{N}\gamma_{xx}(mT_{vz}) = \left(1-\frac{|m|}{N}\right)\gamma_{xx}(mT_{vz})$$

$$E\left[\tilde{r}_{xx2}(mT_{vz})\right] \neq \gamma_{xx}(mT_{vz}) \qquad \lim_{N\to\infty} E\left[\tilde{r}_{xx2}(mT_{vz})\right] = \gamma_{xx}(mT_{vz})$$

$\tilde{r}_{xx2}(mT_{vz})$ je asymptoticky nestranný odhad

# ODHADY AUTOKORELAČNÍ POSLOUPNOSTI

ad2) **rozptyl**

$$\operatorname{var}\left[\tilde{r}_{xx2}(mT_{vz})\right] \approx \frac{1}{N}\sum_{n=-\infty}^{\infty}\left[\left|\gamma_{xx}(nT_{vz})\right|^2 + \gamma_{xx}(nT_{vz} - mT_{vz})\gamma_{xx}(nT_{vz} + mT_{vz})\right]$$

je to menší než pro $\tilde{r}_{xx1}(mT_{vz})$

$$\text{je}-\text{li} \quad \sum_{n=-\infty}^{\infty}\left|\gamma_{xx}(nT_{vz})\right|^2 < \infty, \text{ je} \quad \lim_{N\to\infty} \operatorname{var}\left[\tilde{r}_{xx1}(mT_{vz})\right] = 0$$

a tak $\tilde{r}_{xx2}(mT_{vz})$ je také konzistentní

# ODHADY AUTOKORELAČNÍ POSLOUPNOSTI
## PŘÍKLAD ZE ŽIVOTA