

Plánování Průmyslových Experimentů

Část 2. Metody vyhodnocování experimentů

- Vyhodnocení experimentu: stanovení charakteristik procesu a jeho optimalizace, vyhodnocení vlastností materiálů zpravidla metodami matematické statistiky

Klíčové pojmy: Testování hypotéz, odhady parametrů, regrese, korelace

Plánování Průmyslových Experimentů

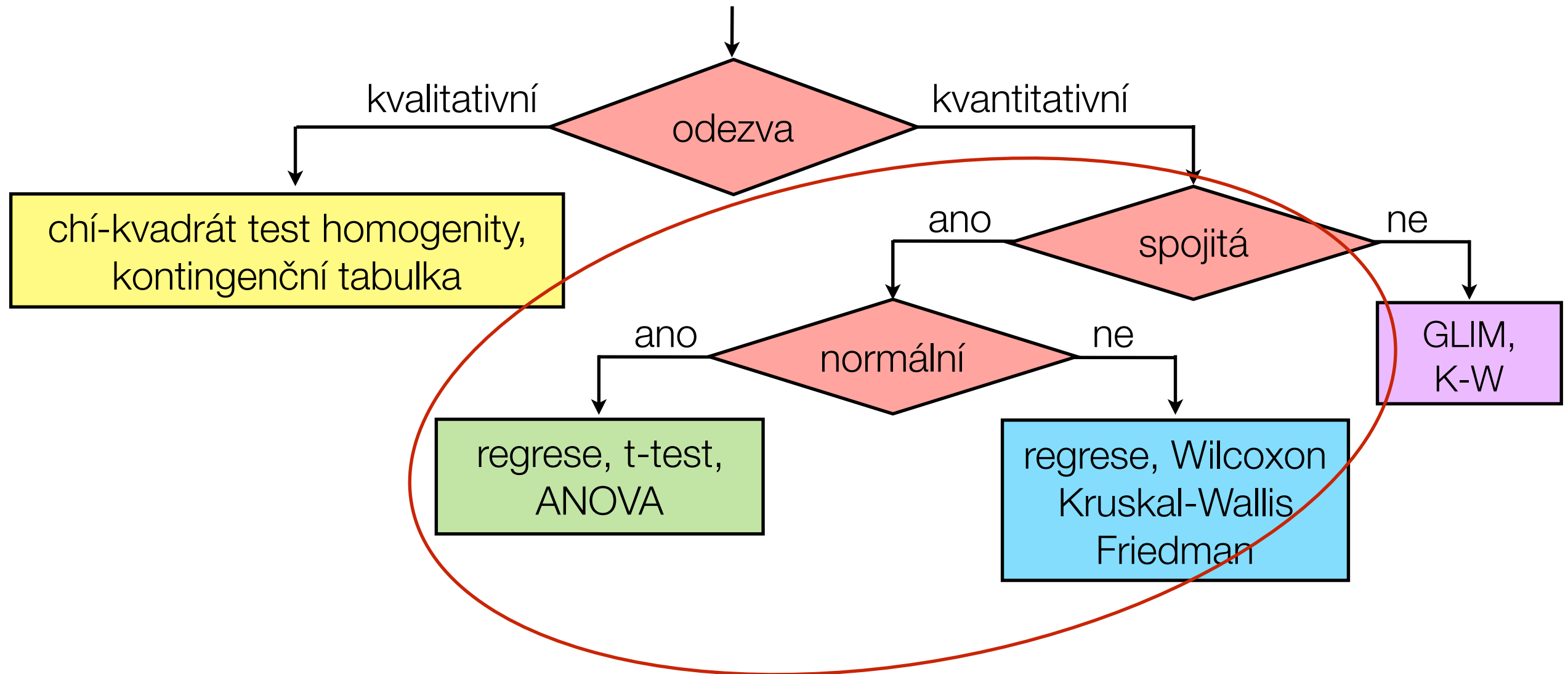
Část 2. Metody vyhodnocování experimentů

- Vyhodnocení experimentu: stanovení charakteristik procesu a jeho optimalizace, vyhodnocení vlastností materiálů zpravidla metodami matematické statistiky

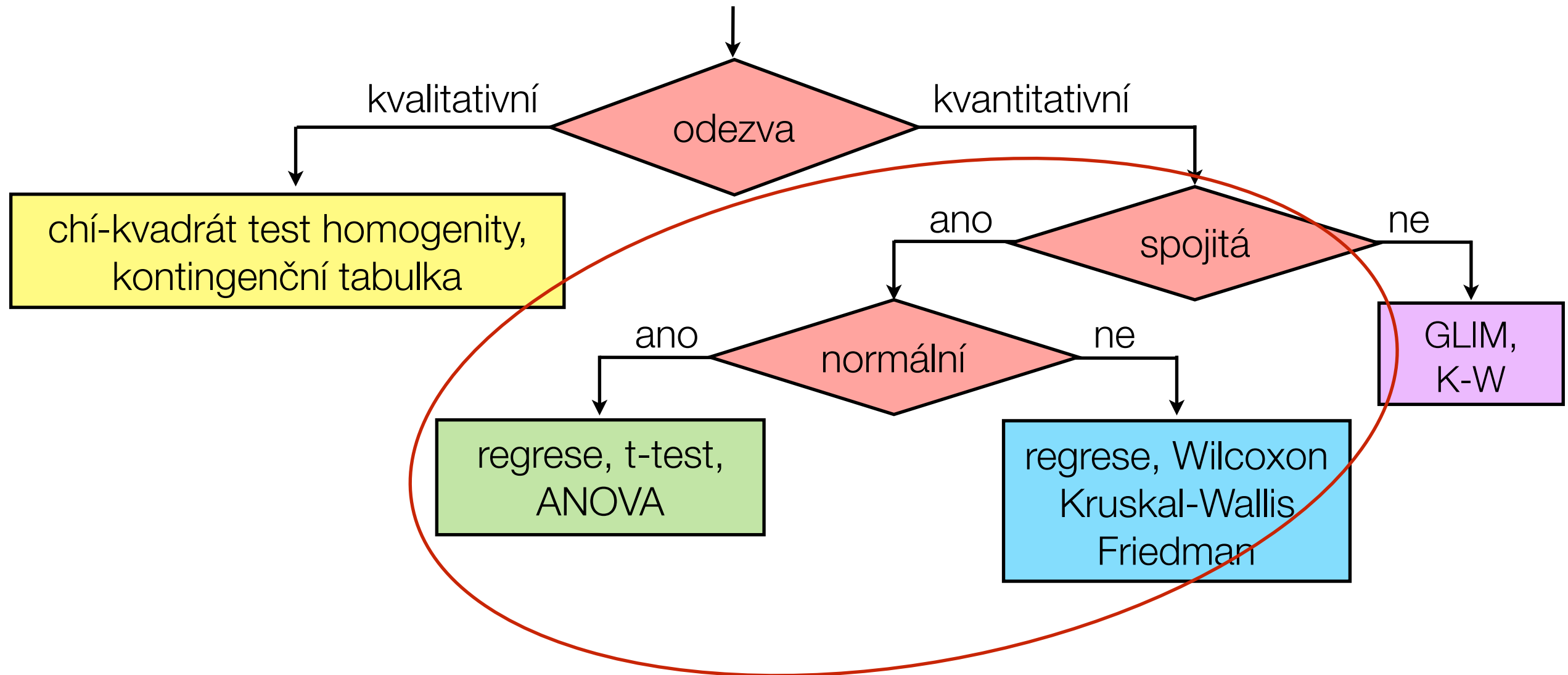
Klíčové pojmy: Testování hypotéz, odhady parametrů, regrese, korelace



2.3. Kvantitativní odezva



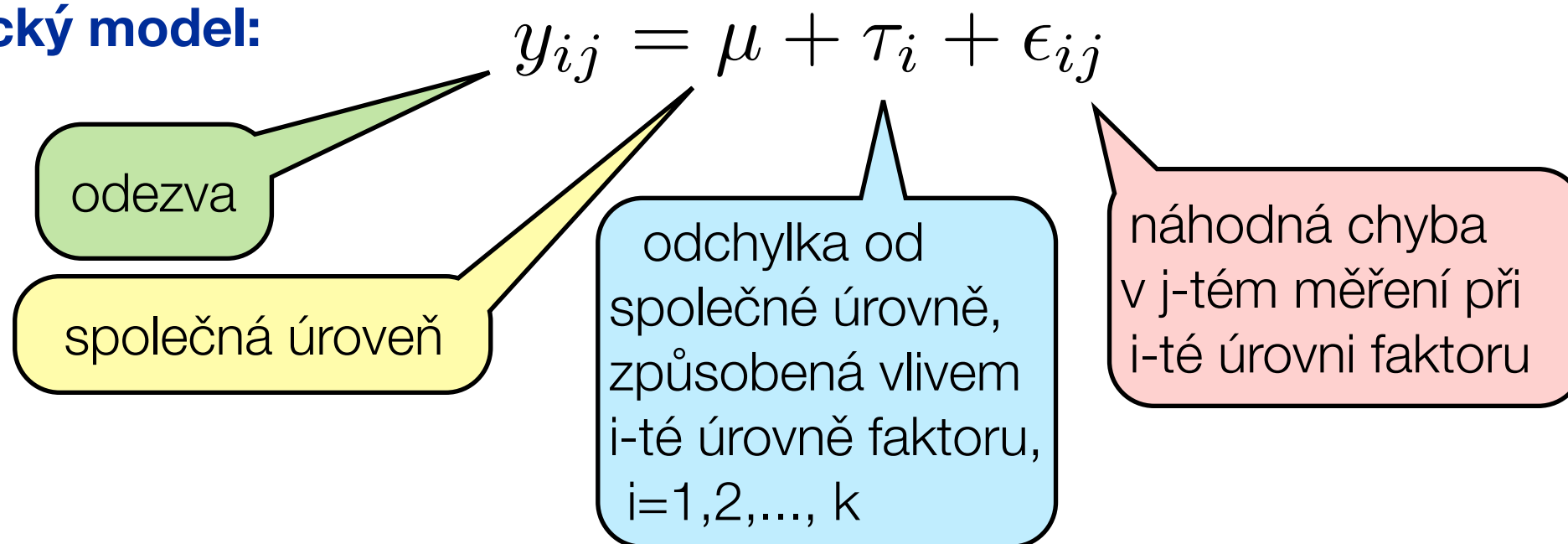
2.3. Kvantitativní odezva



2.3. Kvantitativní odezva, jeden faktor, více úrovní

- uvažujeme vliv pouze jediného (hlavního) faktoru X
- tento faktor může mít k úrovní
- pro každou úroveň provedeme n_i měření (replikací), $i=1,2,\dots,k$
- měření můžeme sdružovat do bloků podle nějakého vedlejšího (blokového) faktoru
- provádíme znáhodnění pořadí měření (buď přes celý experiment nebo pouze uvnitř bloků)

Statistický model:



vliv i-té úrovně faktoru, $i=1,2,\dots, k$: $\mu_i = \mu + \tau_i$

**Budeme testovat hypotézu
proti alternativní hypotéze**

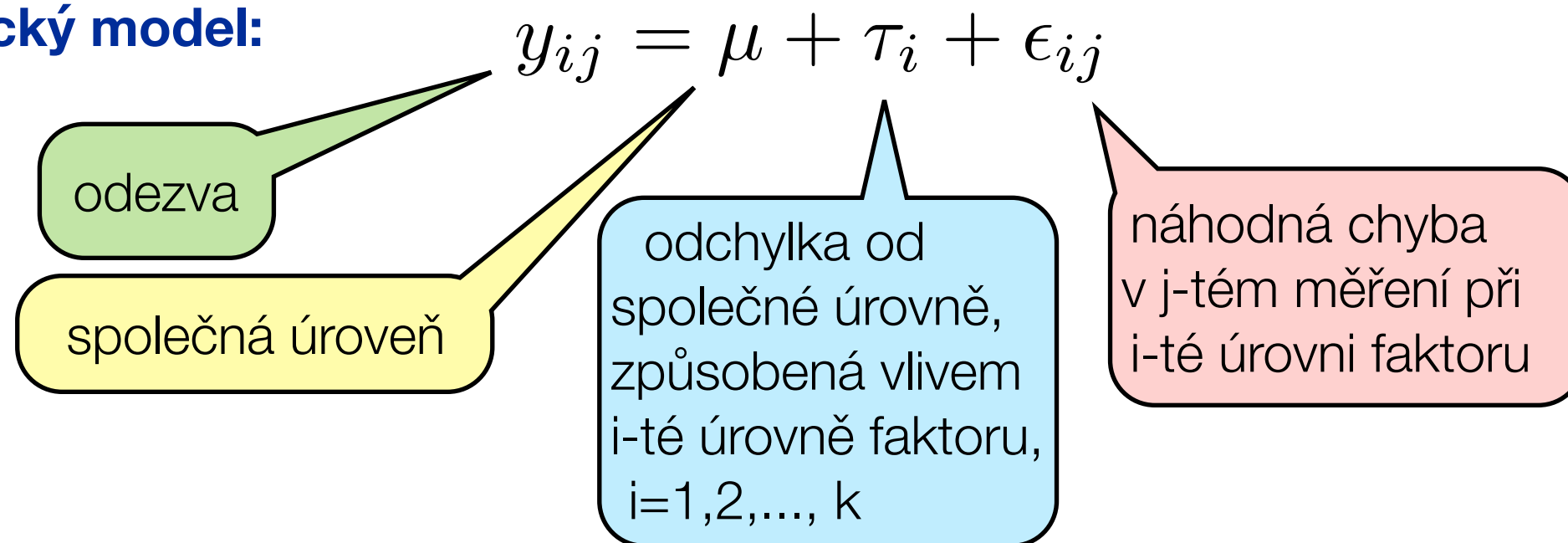
$H_0: \mu_1 = \mu_2 = \dots = \mu_k$

$H_A: \mu_i \neq \mu_j$ pro nějakou dvojici i, j

2.3. Kvantitativní odezva, jeden faktor, více úrovní

- uvažujeme vliv pouze jediného (hlavního) faktoru X
- tento faktor může mít k úrovní
- pro každou úroveň provedeme n_i měření (replikací), $i=1,2,\dots,k$
- měření můžeme sdružovat do bloků podle nějakého vedlejšího (blokového) faktoru
- provádíme znáhodnění pořadí měření (buď přes celý experiment nebo pouze uvnitř bloků)

Statistický model:



vliv i-té úrovně faktoru, $i=1,2,\dots,k$: $\mu_i = \mu + \tau_i$

**Budeme testovat hypotézu
proti alternativní hypotéze**

$H_0: \mu_1 = \mu_2 = \dots = \mu_k$

$H_A: \mu_i \neq \mu_j$ pro nějakou dvojici i,j



2.3. Kvantitativní odezva, jeden faktor, více úrovní

pořadové číslo měření (replikace)

úroveň faktoru (ošetření)

	1	2	3	...	n
1	y_{11}	y_{12}	y_{13}	...	y_{1n_1}
2	y_{21}	y_{22}	y_{23}	...	y_{2n_2}
3	y_{31}	y_{32}	y_{33}	...	y_{3n_3}
:	:	:	:	$\ddots$	:
k	y_{k1}	y_{k2}	y_{k3}	...	y_{kn_k}

2.3. Kvantitativní odezva, jeden faktor, více úrovní

pořadové číslo měření (replikace)

úroveň faktoru (ošetření)

	1	2	3	...	n
1	y_{11}	y_{12}	y_{13}	...	y_{1n_1}
2	y_{21}	y_{22}	y_{23}	...	y_{2n_2}
3	y_{31}	y_{32}	y_{33}	...	y_{3n_3}
:	:	:	:	$\ddots$	:
k	y_{k1}	y_{k2}	y_{k3}	...	y_{kn_k}



2.3. Kvantitativní odezva, jeden faktor o více úrovních

Srovnání pevnosti vláken od tří dodavatelů

Chceme porovnat kvalitu vláken dodávaných třemi různými výrobci. Jako kritérium kvality je zvolena pevnost vláken v tahu.

2.3. Kvantitativní odezva, jeden faktor o více úrovních

Srovnání pevnosti vláken od tří dodavatelů

Chceme porovnat kvalitu vláken dodávaných třemi různými výrobci. Jako kritérium kvality je zvolena pevnost vláken v tahu.

- Odezva: pevnost vlákna
- Faktor: dodavatel (3 úrovně: dodavatel 1, dodavatel 2, dodavatel 3)

2.3. Kvantitativní odezva, jeden faktor o více úrovních

Srovnání pevnosti vláken od tří dodavatelů

Chceme porovnat kvalitu vláken dodávaných třemi různými výrobci. Jako kritérium kvality je zvolena pevnost vláken v tahu.

- Odezva: pevnost vlákna
- Faktor: dodavatel (3 úrovně: dodavatel 1, dodavatel 2, dodavatel 3)

Jedná se o jedno-faktorový experiment se 3 úrovněmi faktoru, s nezávislými měřeními
=> k vyhodnocení použijeme metodu analýzy rozptylu ANOVA pro jednoduché třídění.

2.3. Kvantitativní odezva, jeden faktor o více úrovních

Srovnání pevnosti vláken od tří dodavatelů

Chceme porovnat kvalitu vláken dodávaných třemi různými výrobci. Jako kritérium kvality je zvolena pevnost vláken v tahu.

- Odezva: pevnost vlákna
- Faktor: dodavatel (3 úrovně: dodavatel 1, dodavatel 2, dodavatel 3)

Jedná se o jedno-faktorový experiment se 3 úrovněmi faktoru, s nezávislými měřeními
=> k vyhodnocení použijeme metodu analýzy rozptylu ANOVA pro jednoduché třídění.

nulová hypotéza $H_0: \mu_1 = \mu_2 = \mu_3$

alternativní hypotéza $H_A: \mu_i \neq \mu_j$ pro některou dvojici i, j (oboustranná)

hladina významnosti $\alpha = 5\%$

2.3. Kvantitativní odezva, jeden faktor o více úrovních

Srovnání pevnosti vláken od tří dodavatelů

Chceme porovnat kvalitu vláken dodávaných třemi různými výrobci. Jako kritérium kvality je zvolena pevnost vláken v tahu.

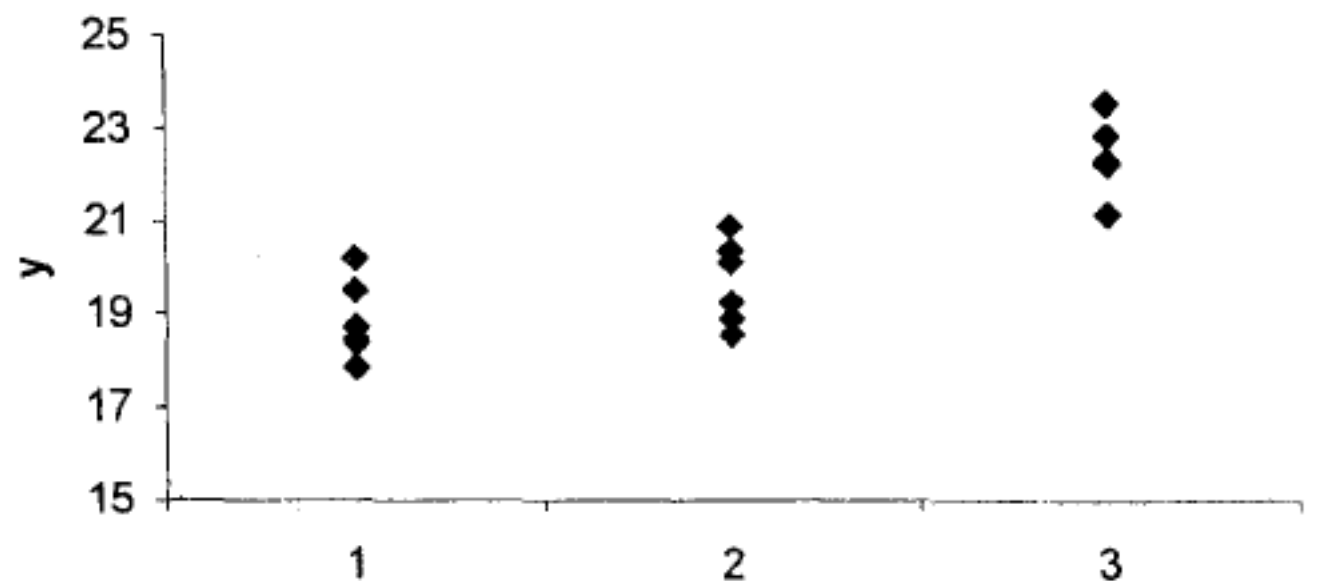
- Odezva: pevnost vlákna
- Faktor: dodavatel (3 úrovně: dodavatel 1, dodavatel 2, dodavatel 3)

Jedná se o jedno-faktorový experiment se 3 úrovněmi faktoru, s nezávislými měřeními
=> k vyhodnocení použijeme metodu analýzy rozptylu ANOVA pro jednoduché třídění.

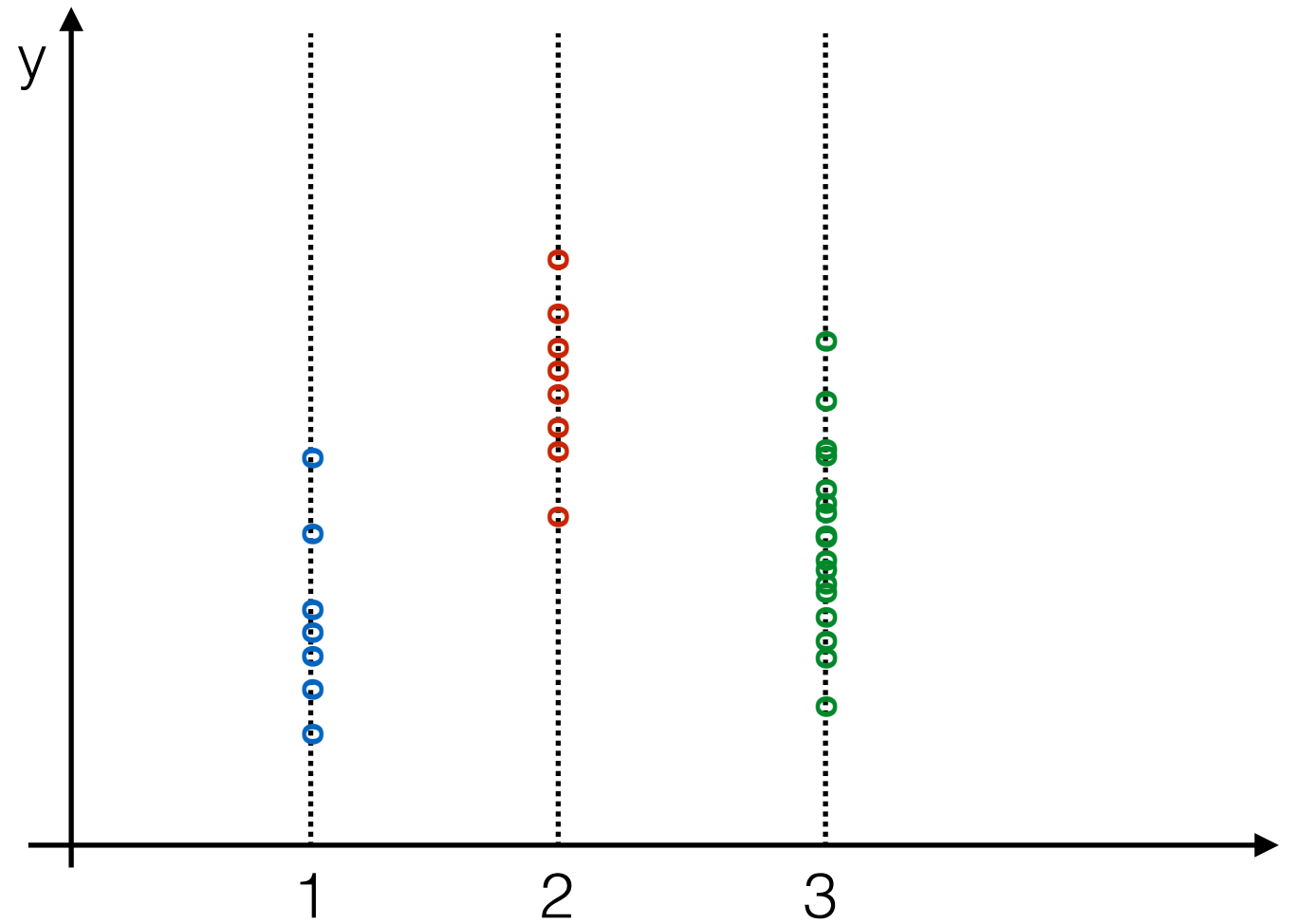
nulová hypotéza $H_0: \mu_1 = \mu_2 = \mu_3$

alternativní hypotéza $H_A: \mu_i \neq \mu_j$ pro některou dvojici i, j (oboustranná)

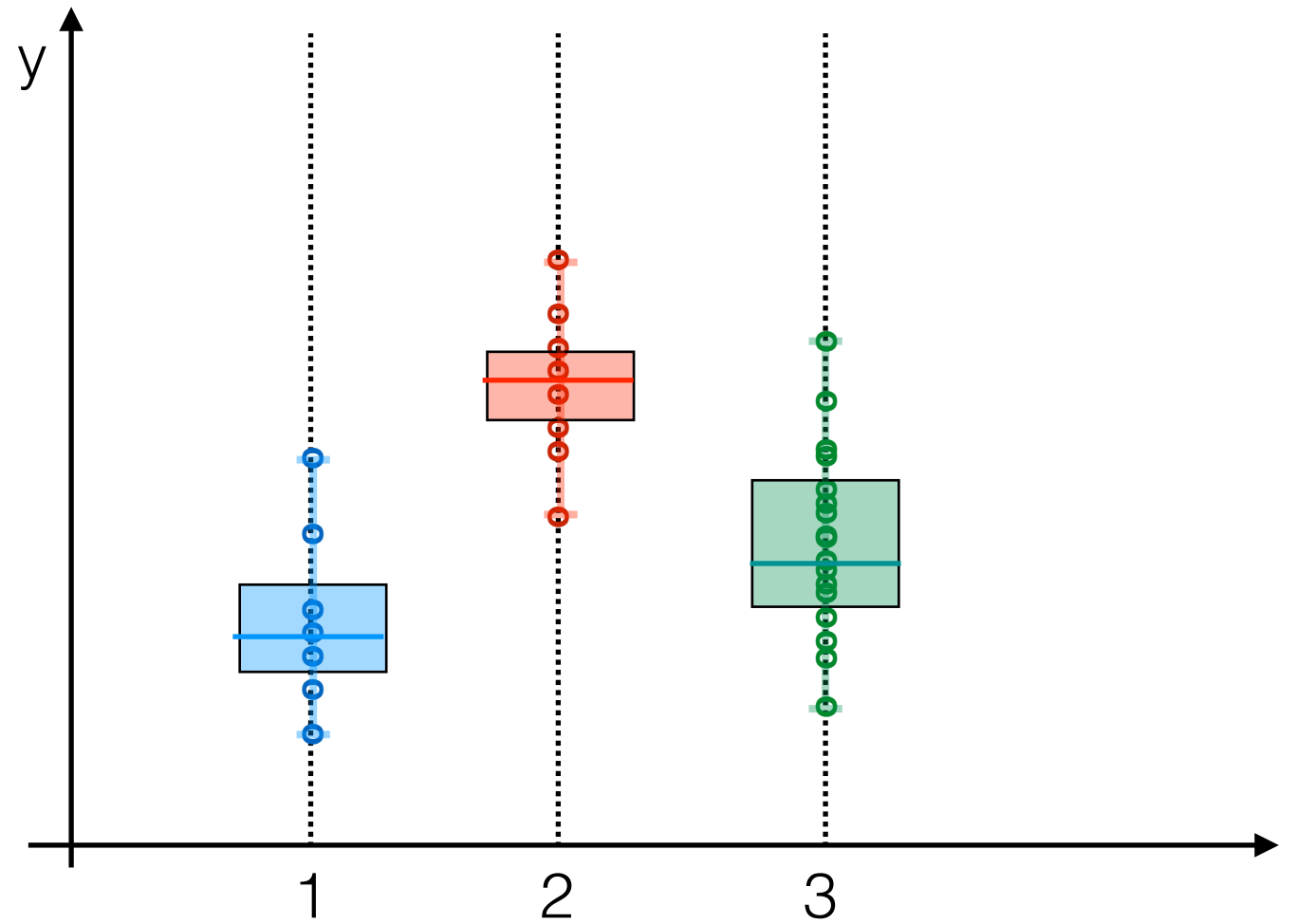
hladina významnosti $\alpha = 5\%$



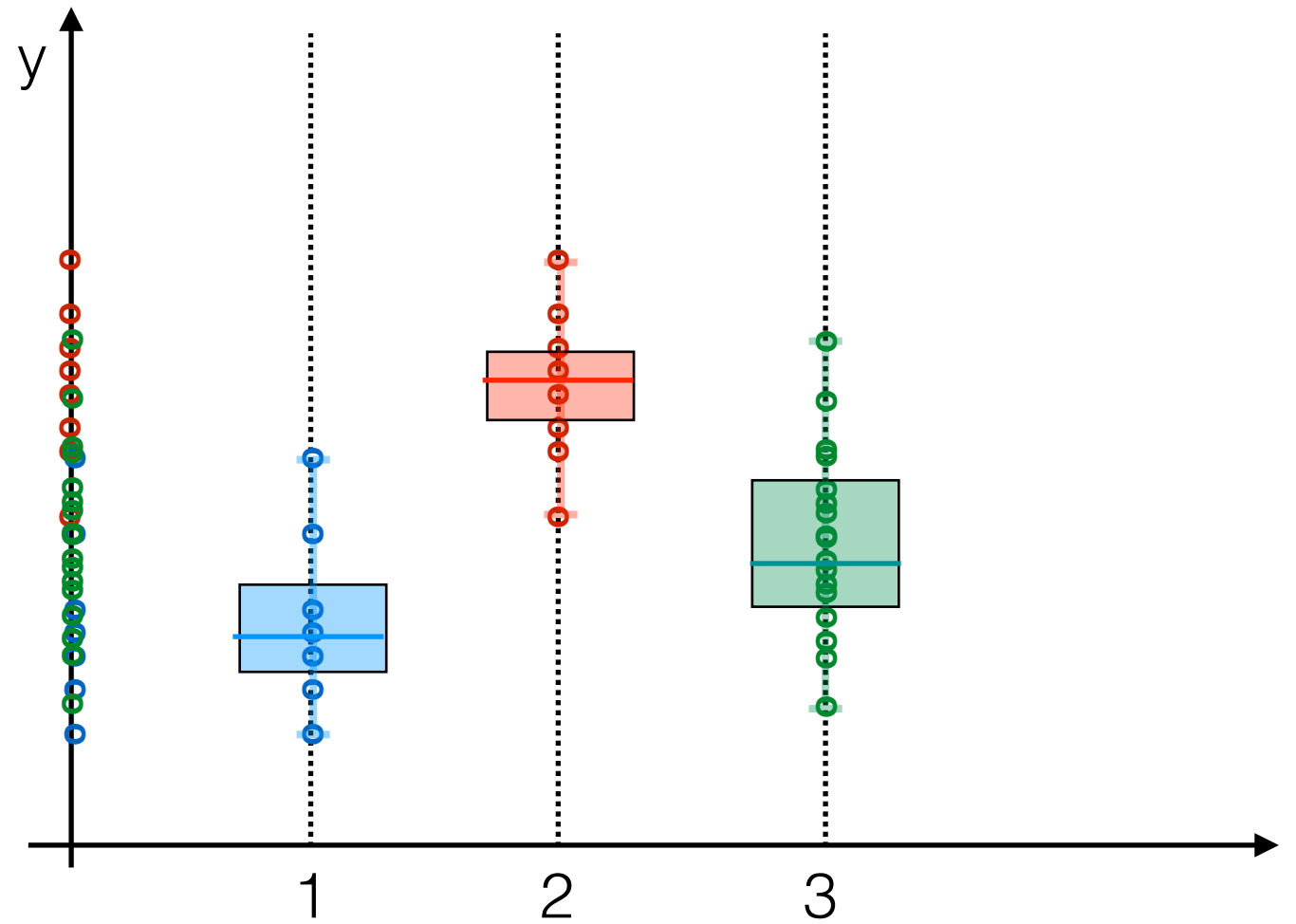
2.3.1. Analýza rozptylu



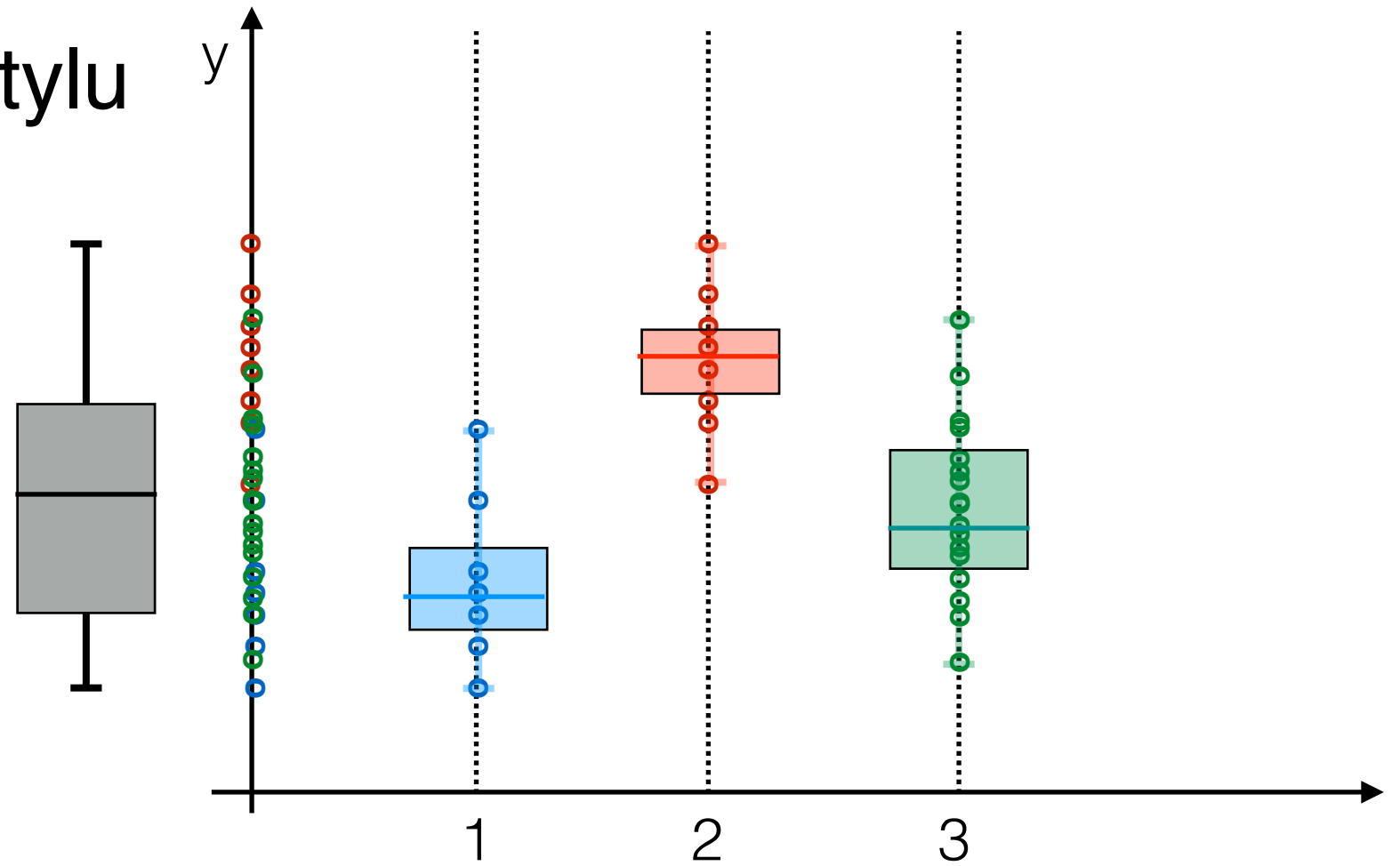
2.3.1. Analýza rozptylu



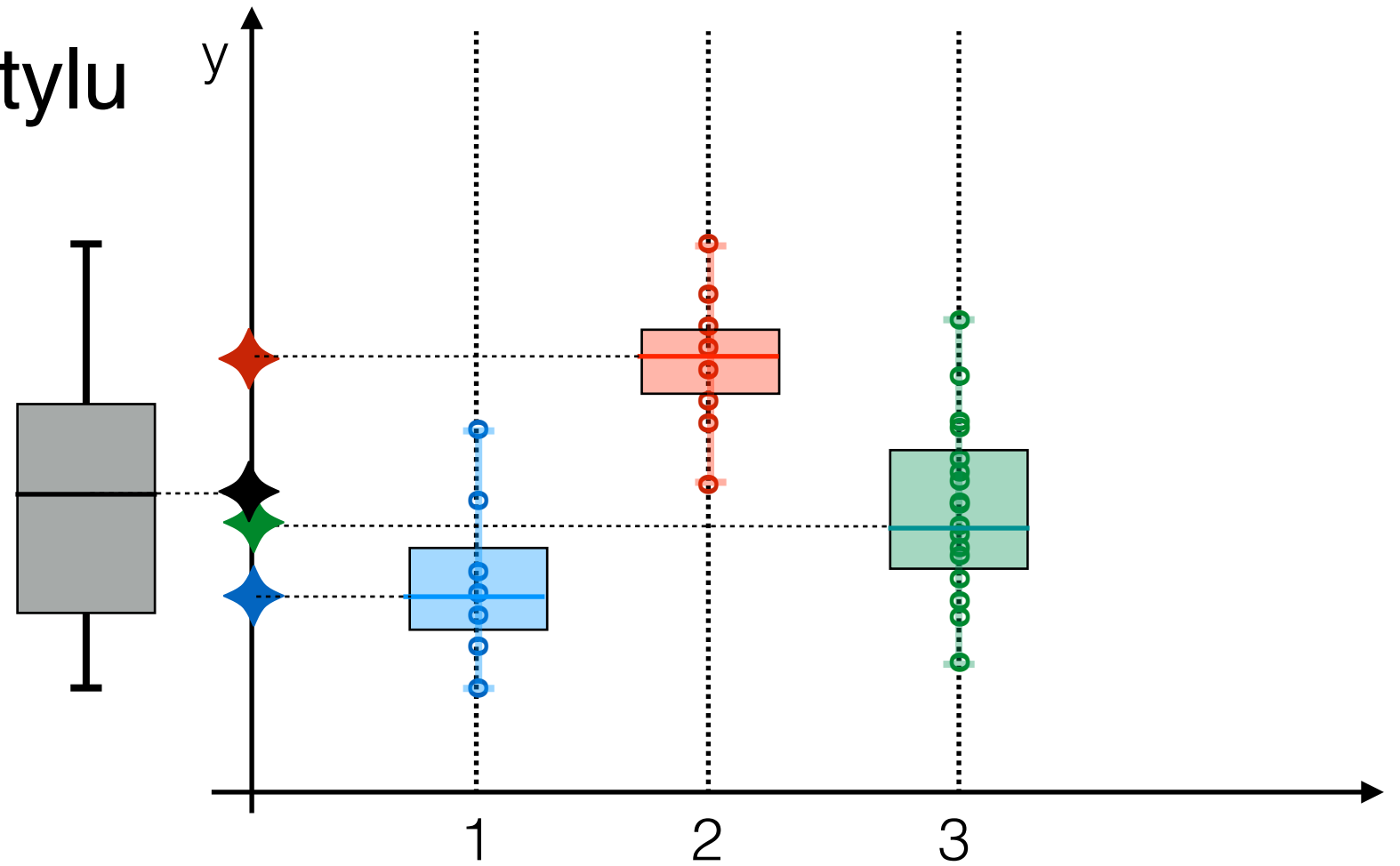
2.3.1. Analýza rozptylu



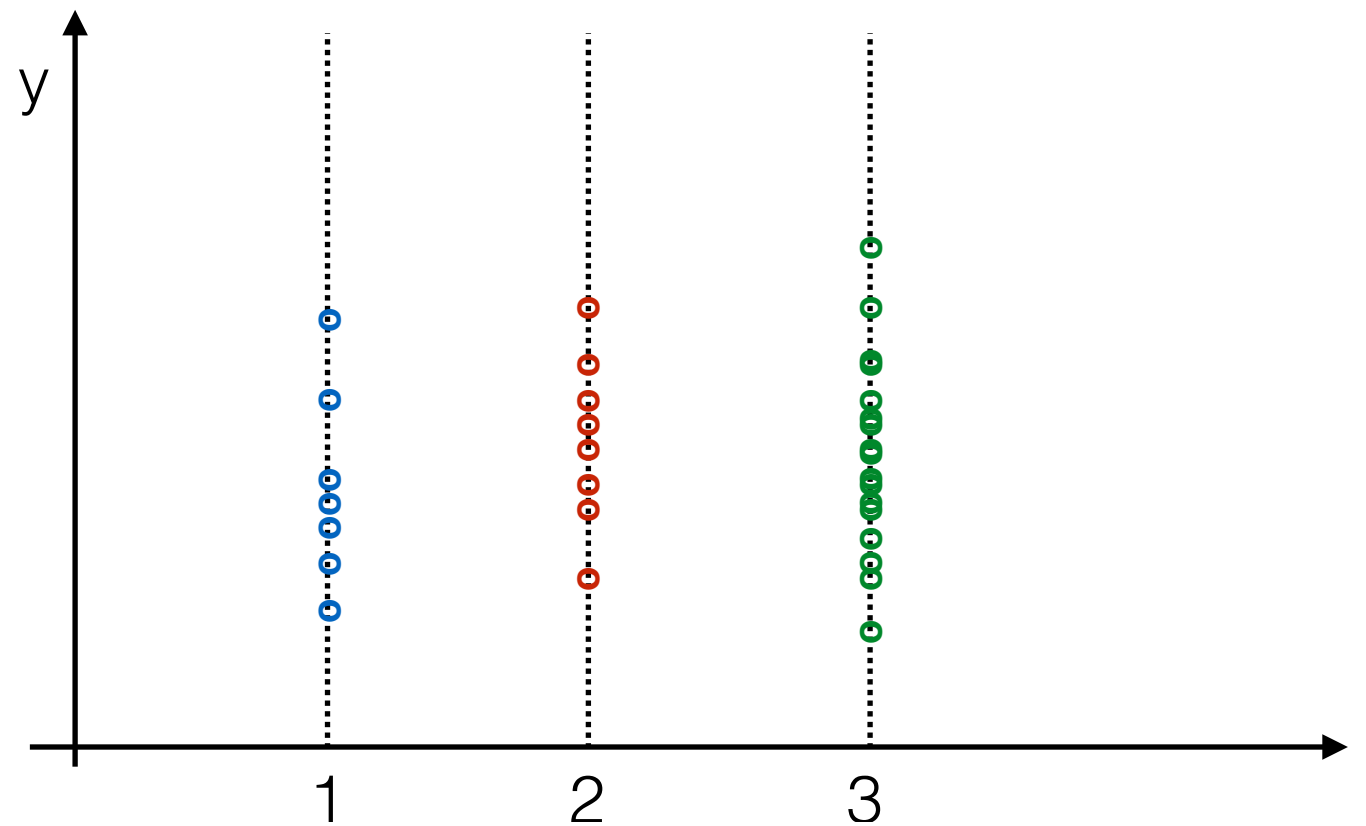
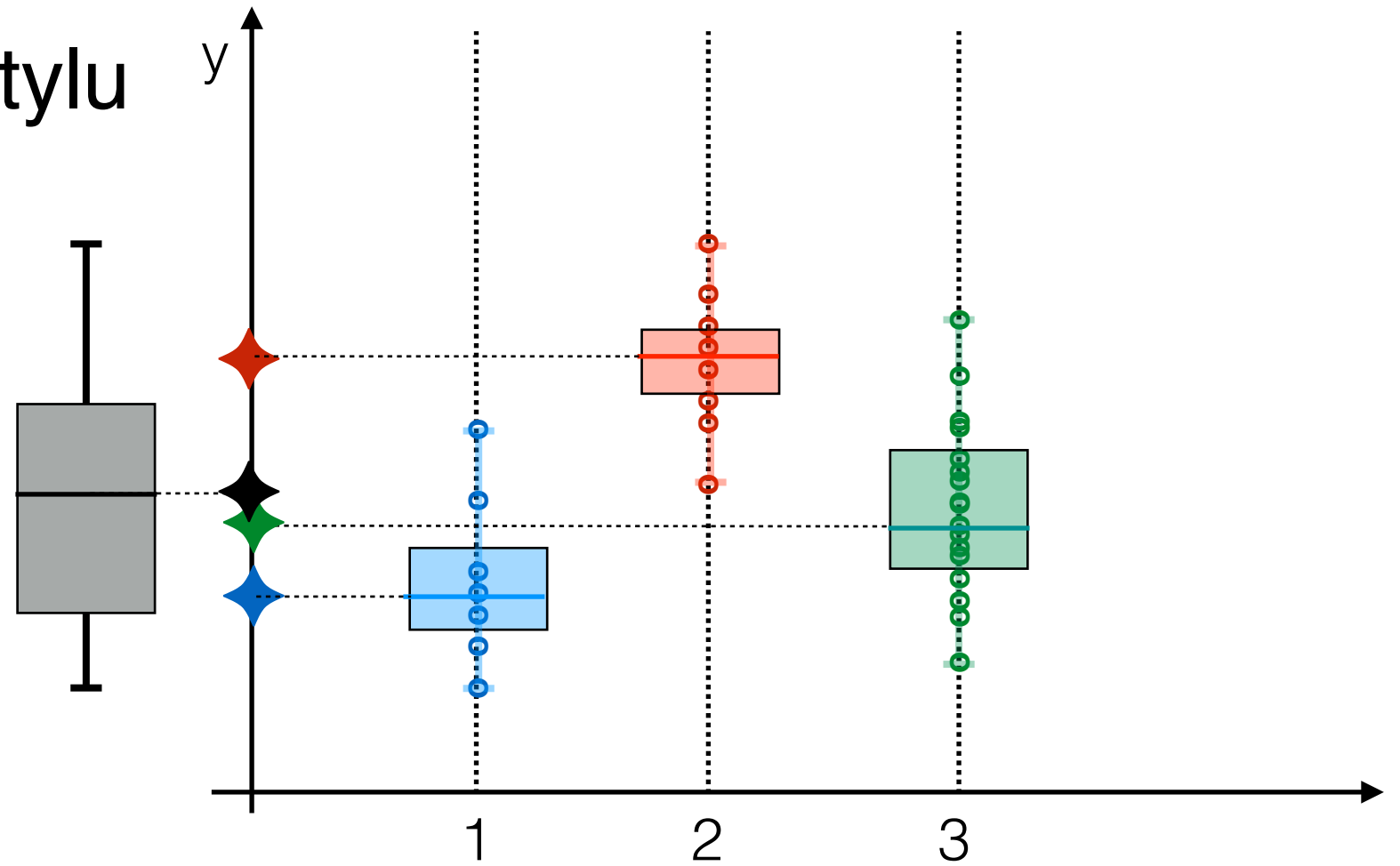
2.3.1. Analýza rozptylu



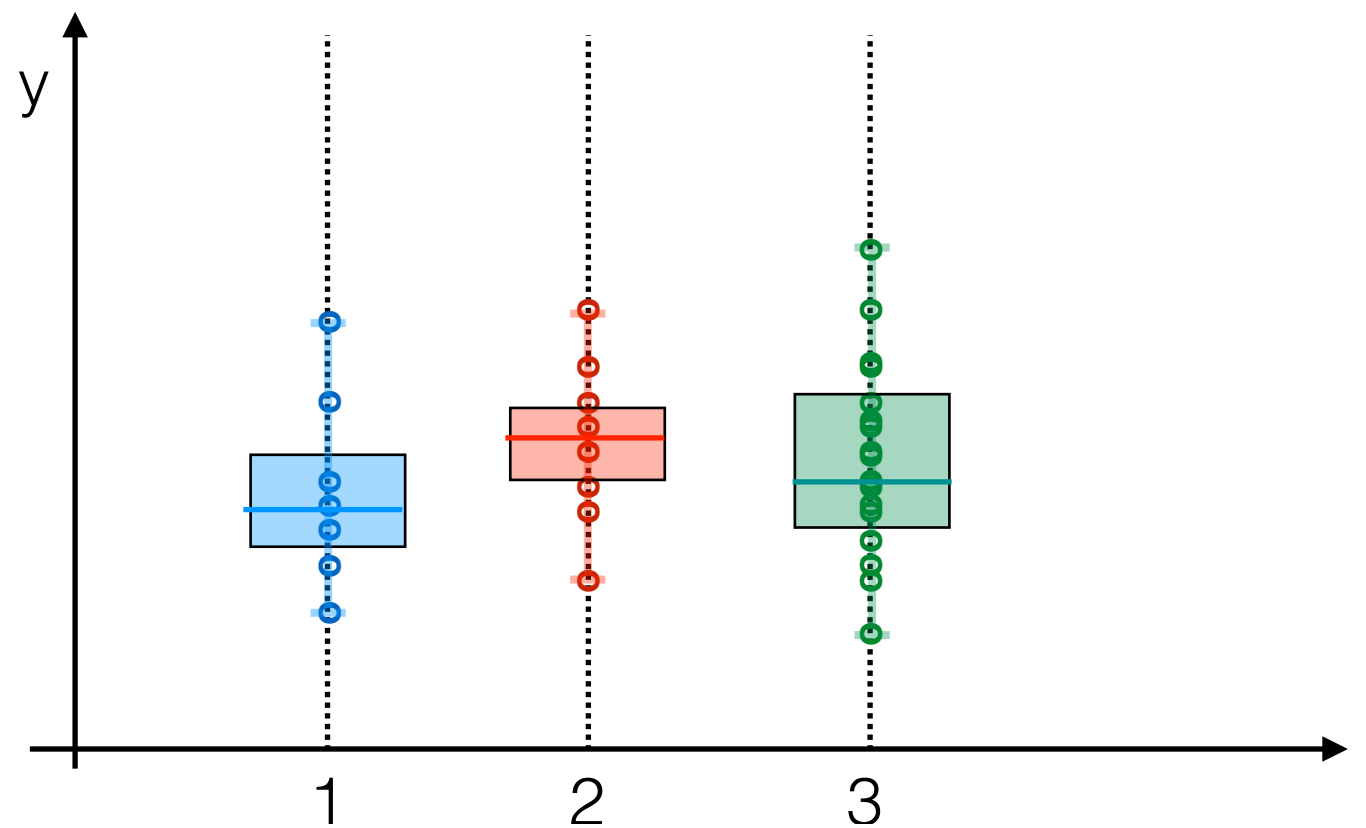
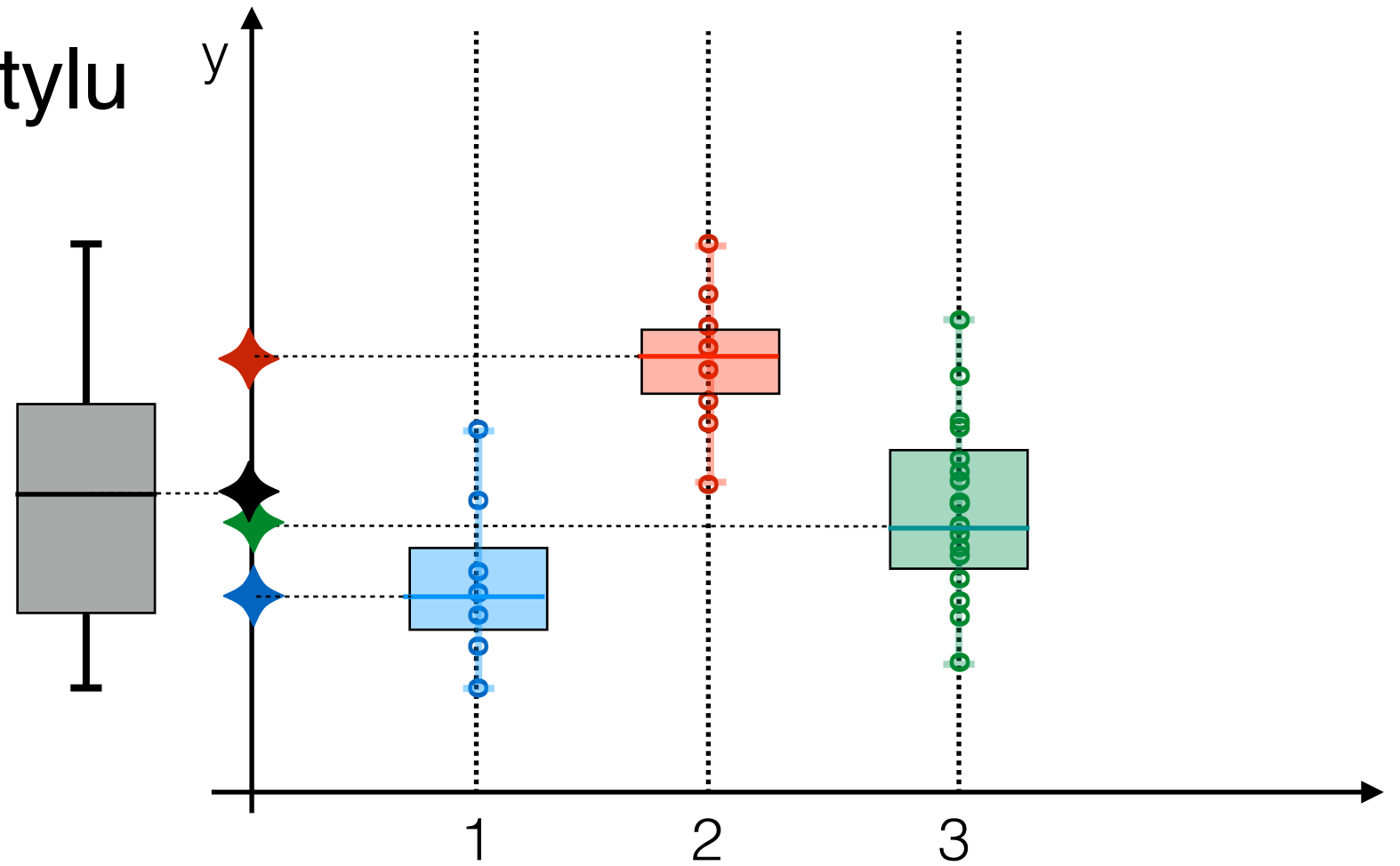
2.3.1. Analýza rozptylu



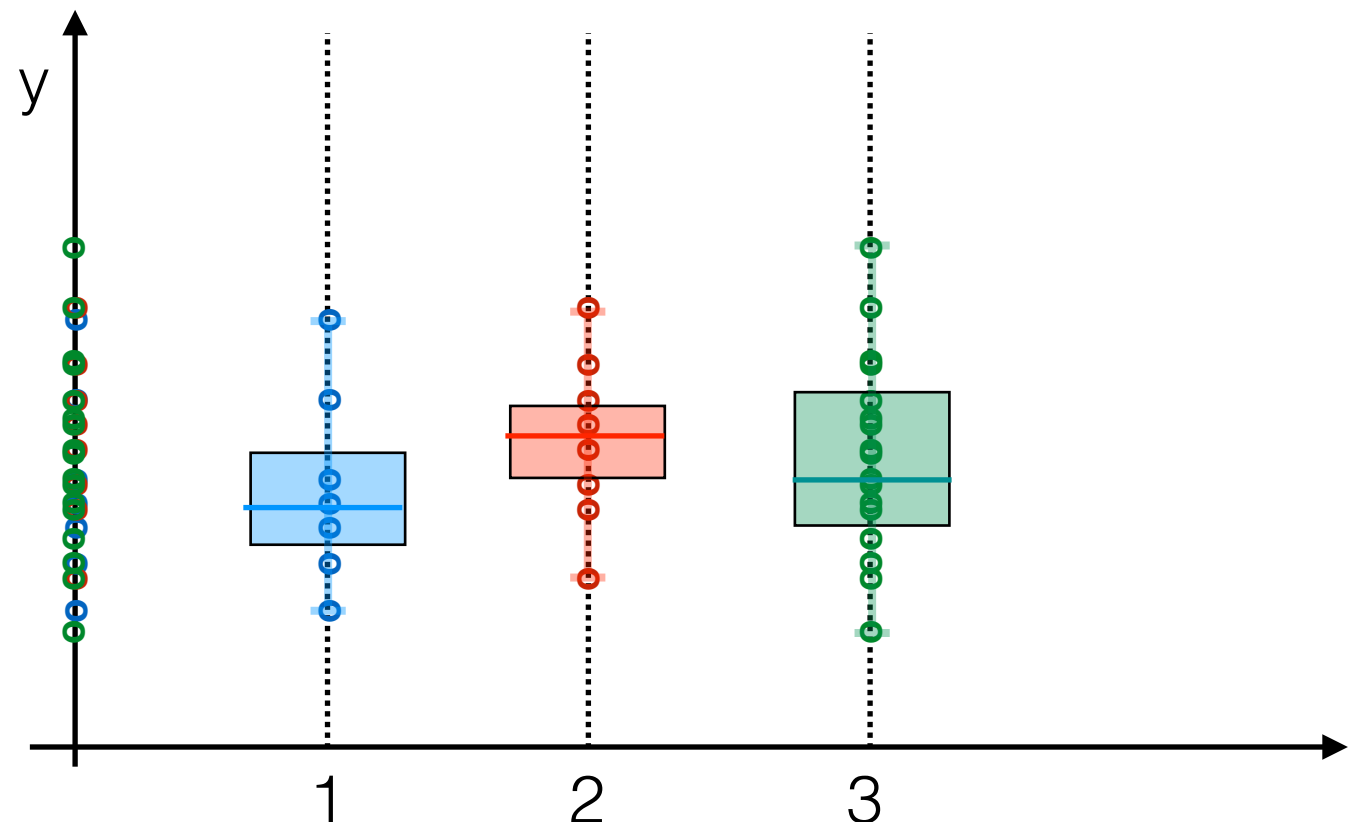
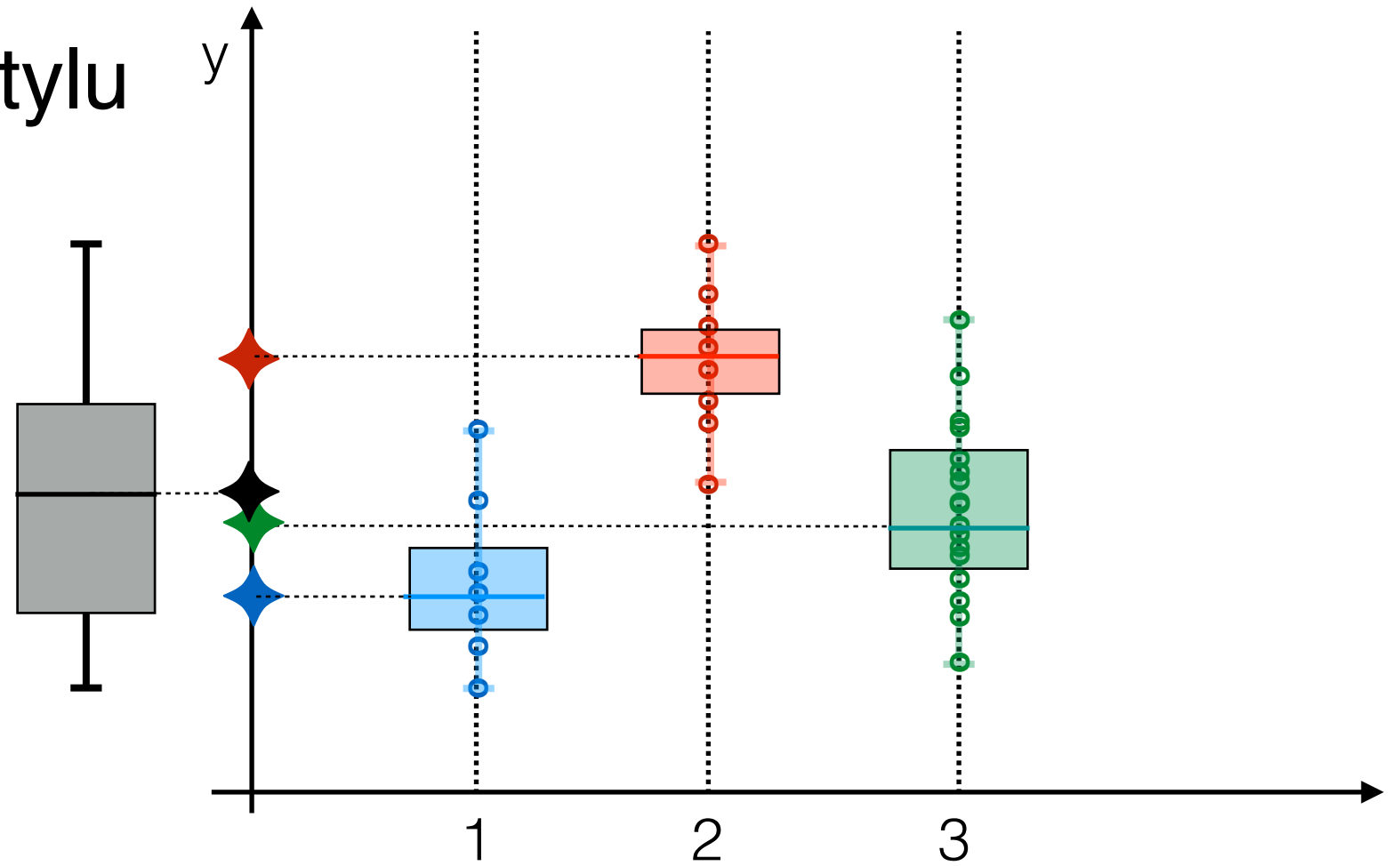
2.3.1. Analýza rozptylu



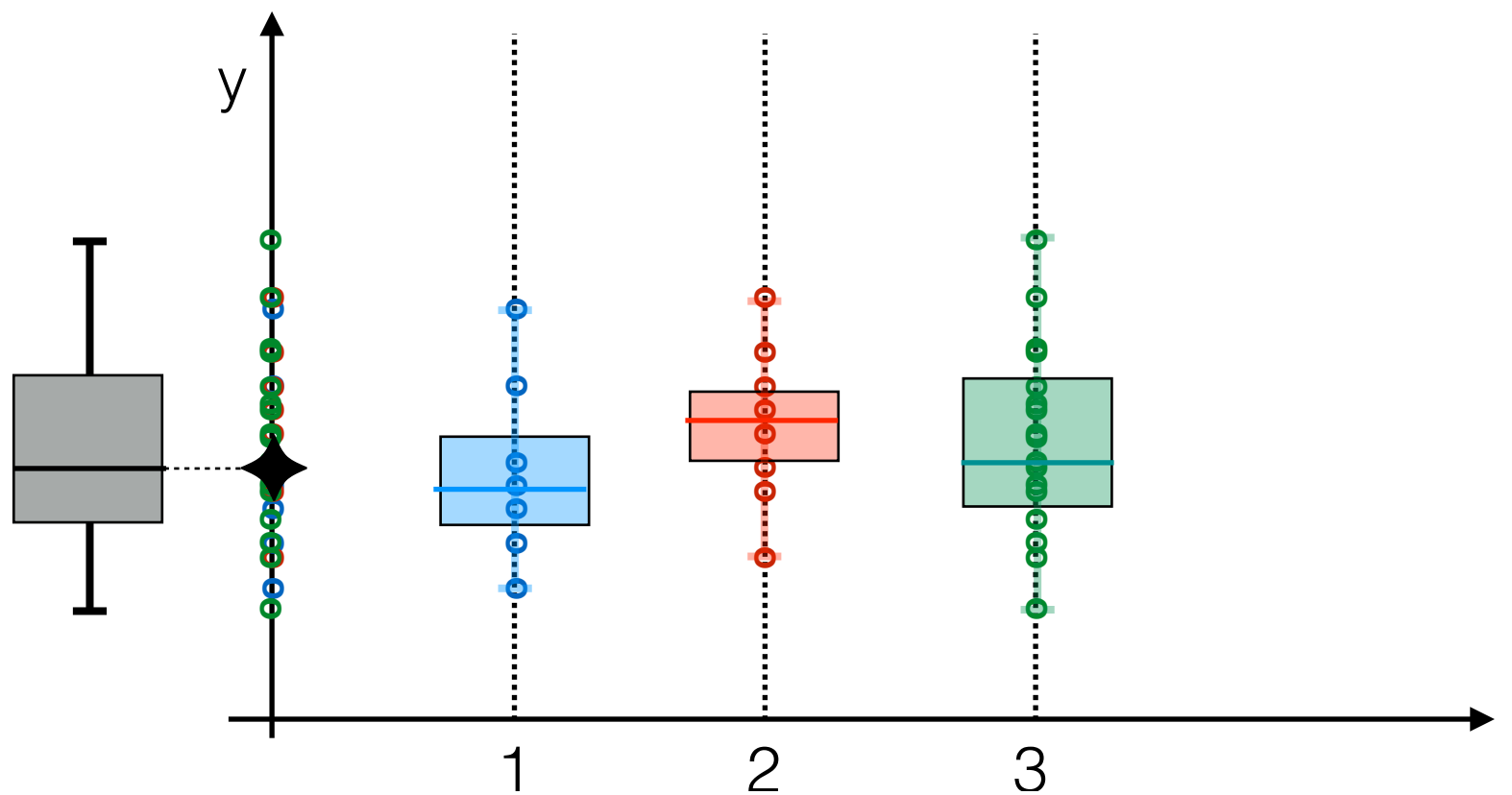
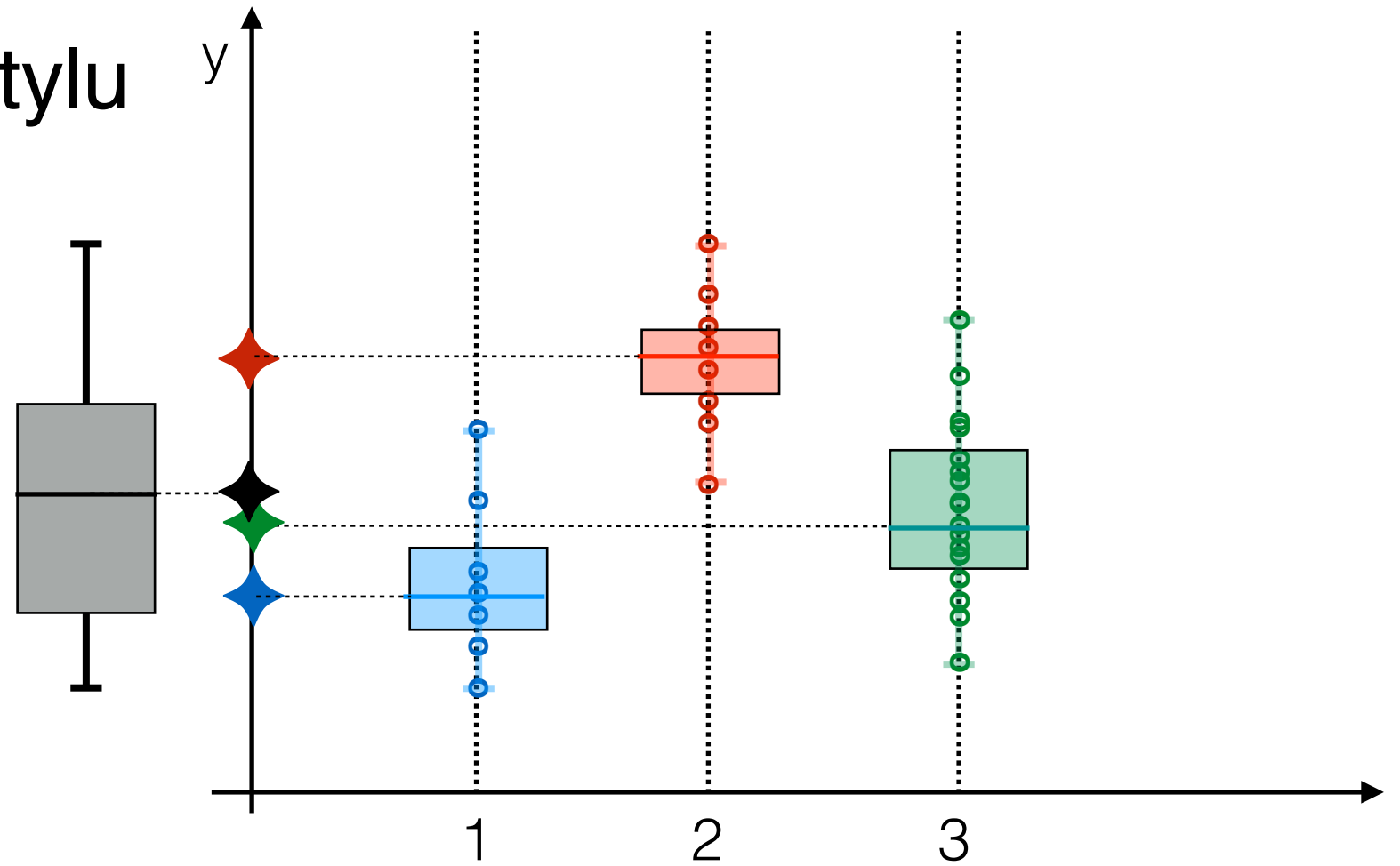
2.3.1. Analýza rozptylu



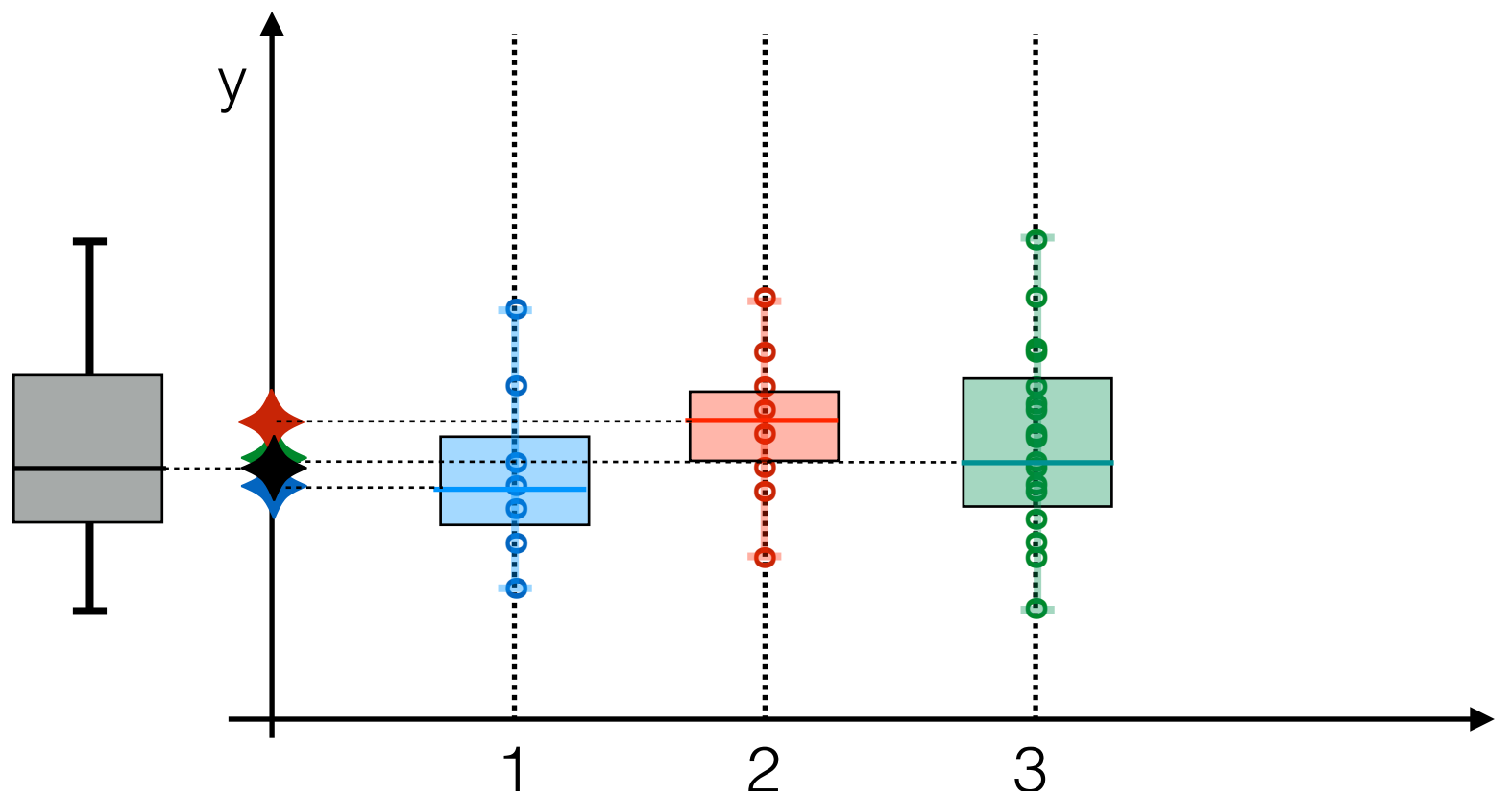
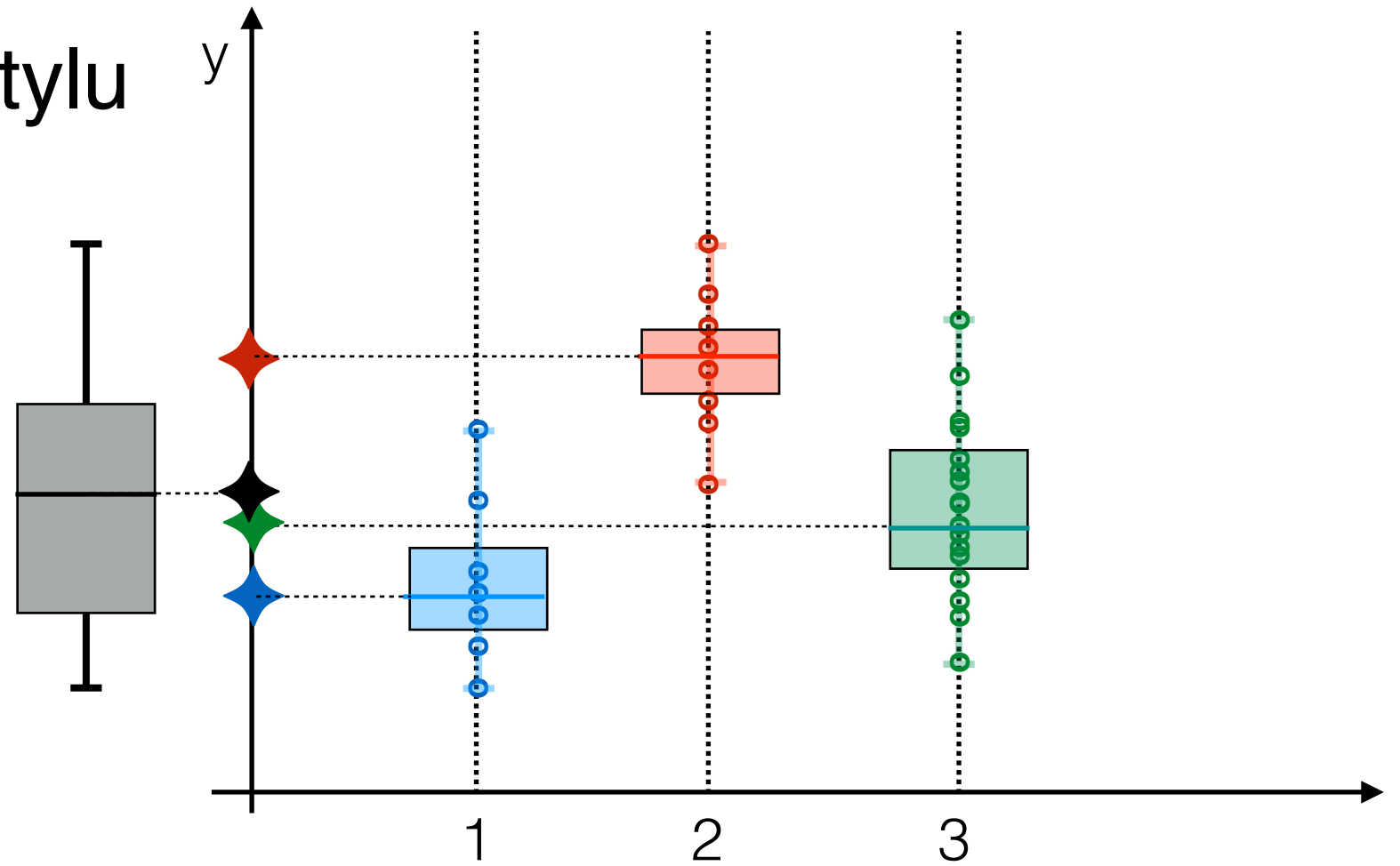
2.3.1. Analýza rozptylu



2.3.1. Analýza rozptylu

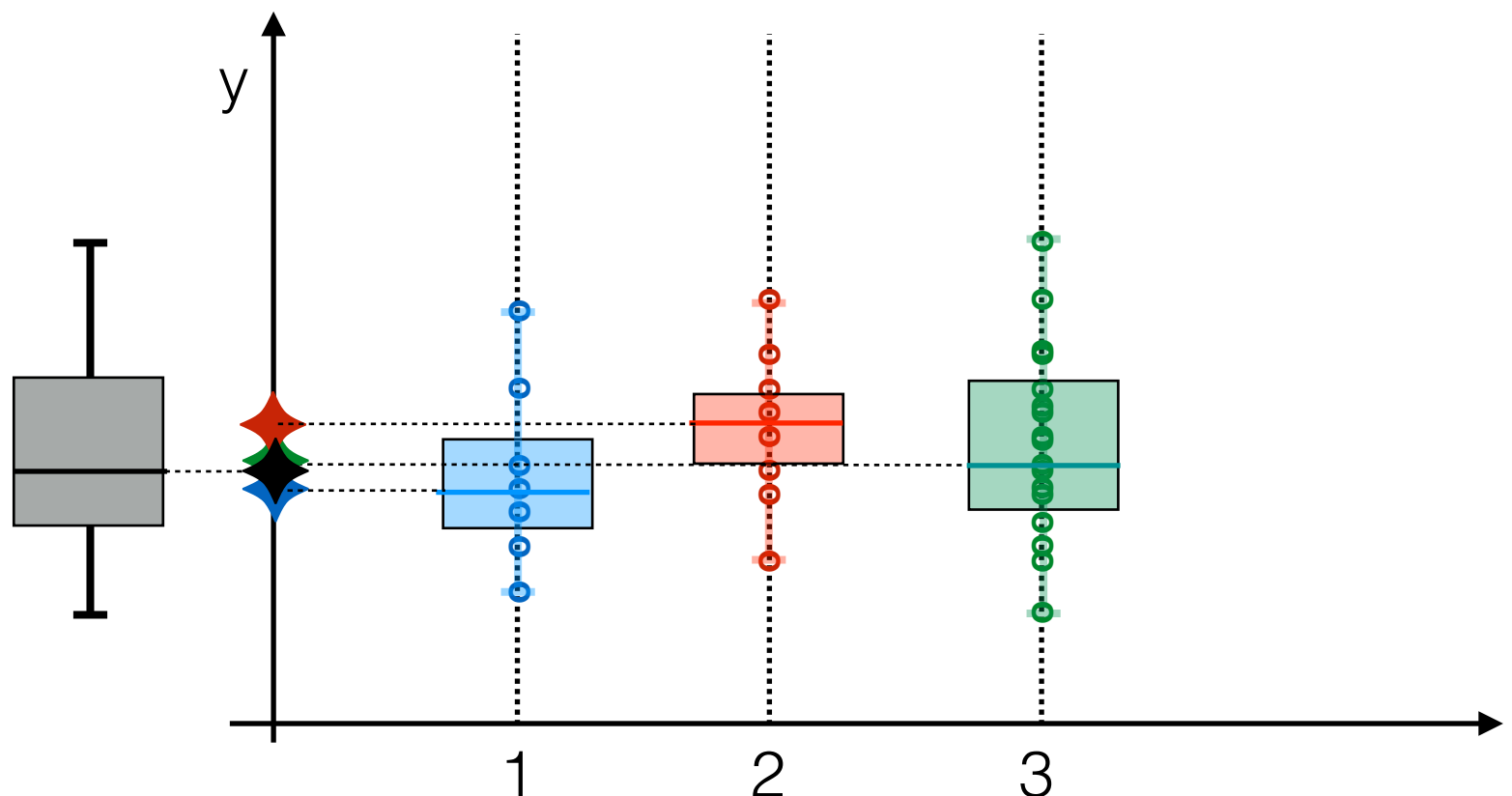
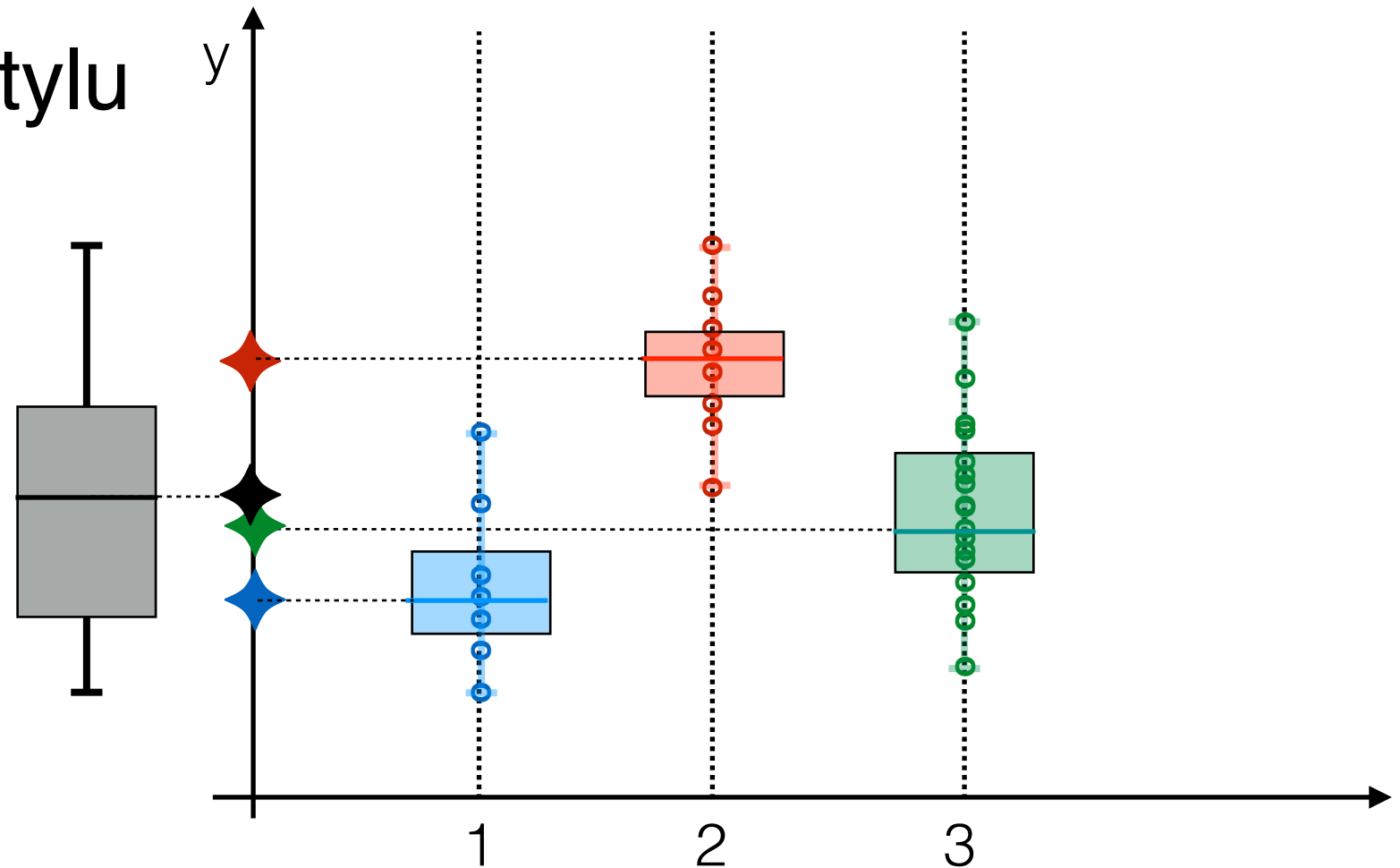


2.3.1. Analýza rozptylu



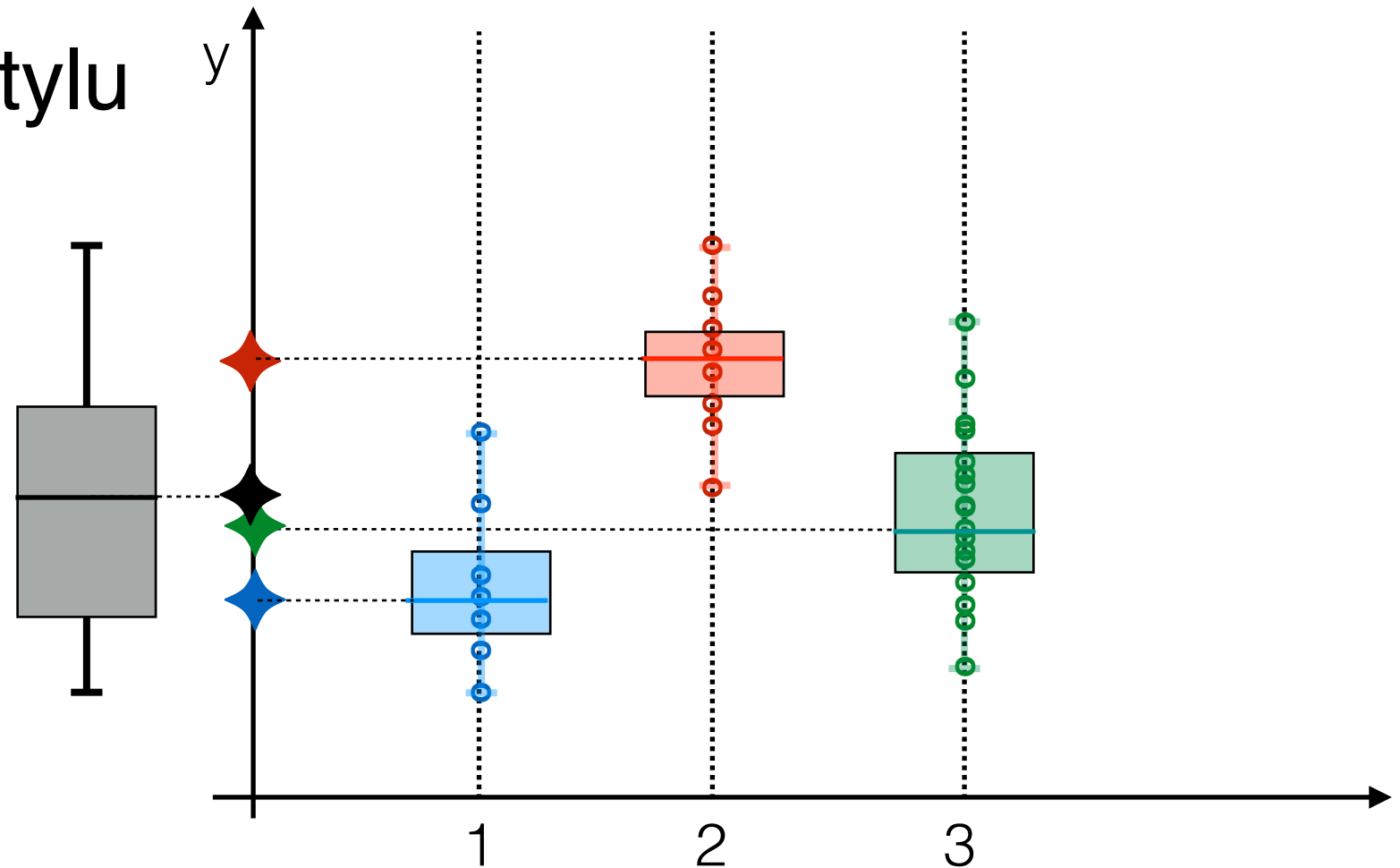
2.3.1. Analýza rozptylu

Rozptyl mezi skupinami je **velký** vzhledem k součtu rozptylů uvnitř skupin => střední hodnoty **jsou různé** => nulovou hypotézu **zamítáme**

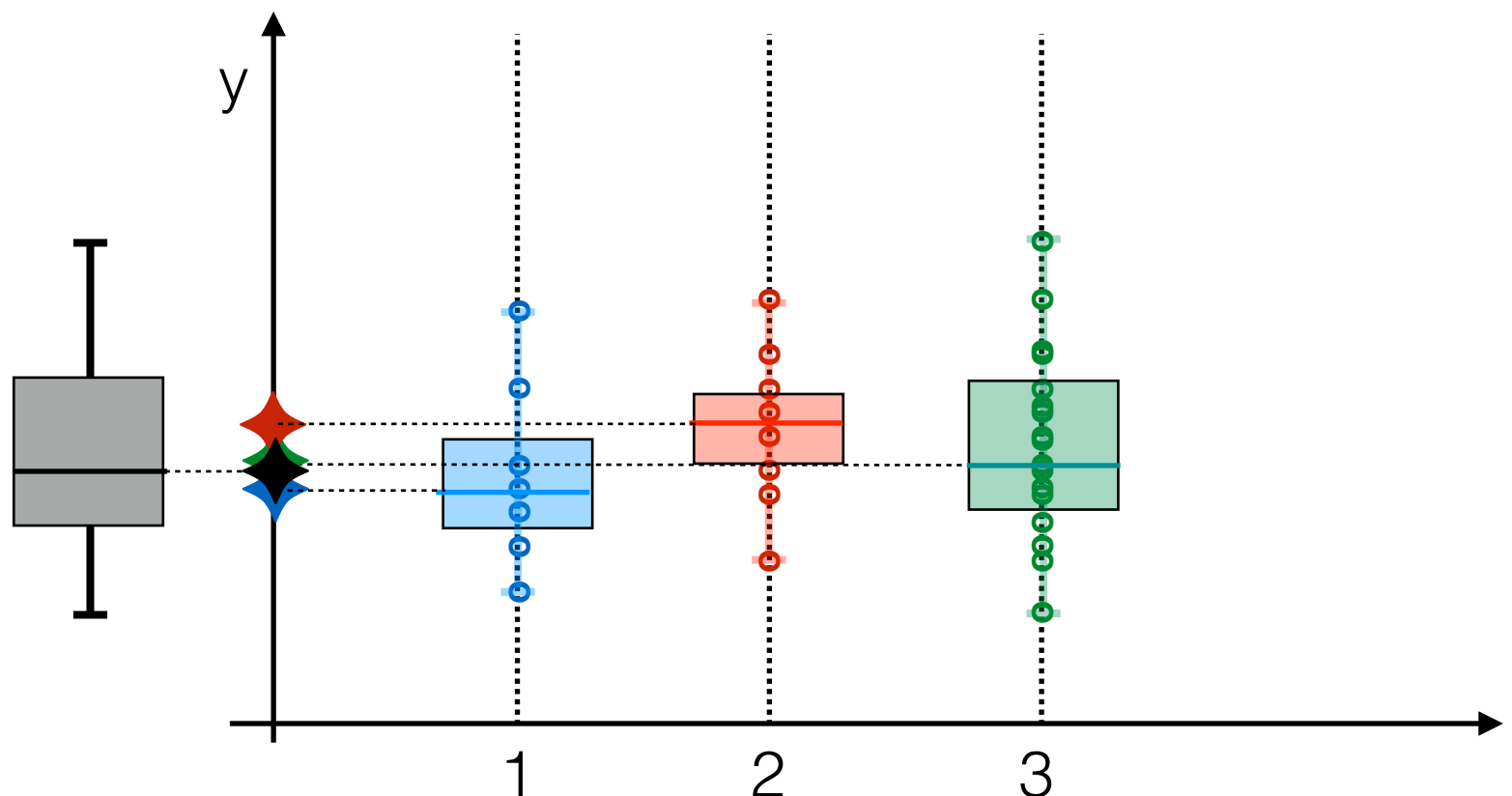


2.3.1. Analýza rozptylu

Rozptyl mezi skupinami je **velký** vzhledem k součtu rozptylů uvnitř skupin => střední hodnoty **jsou různé** => nulovou hypotézu **zamítáme**



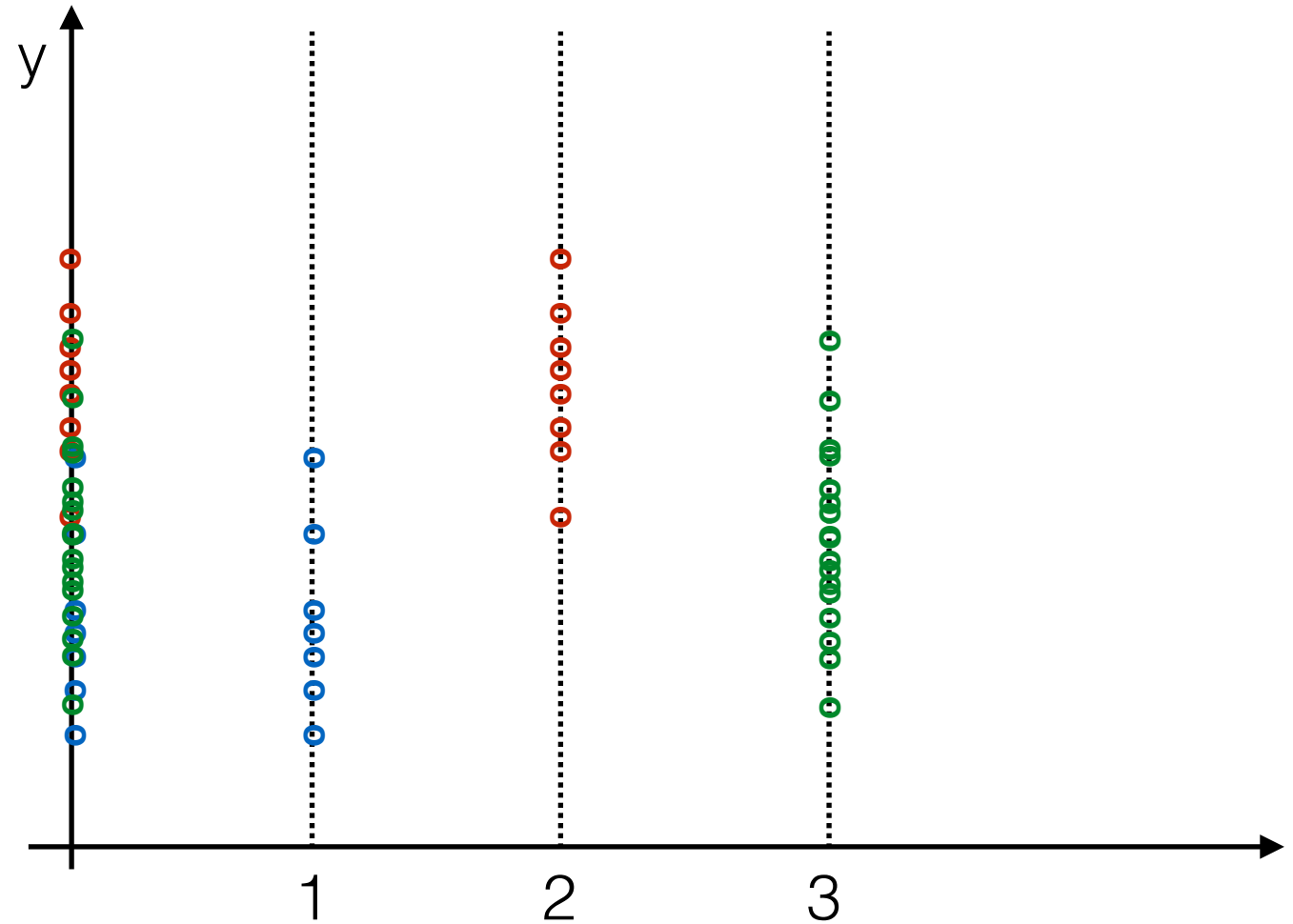
Rozptyl mezi skupinami je **malý** vzhledem k součtu rozptylů uvnitř skupin => střední hodnoty lze považovat **za nerozlišitelné** => nulovou hypotézu **nelze zamítnout**



2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovně jednoho faktoru)



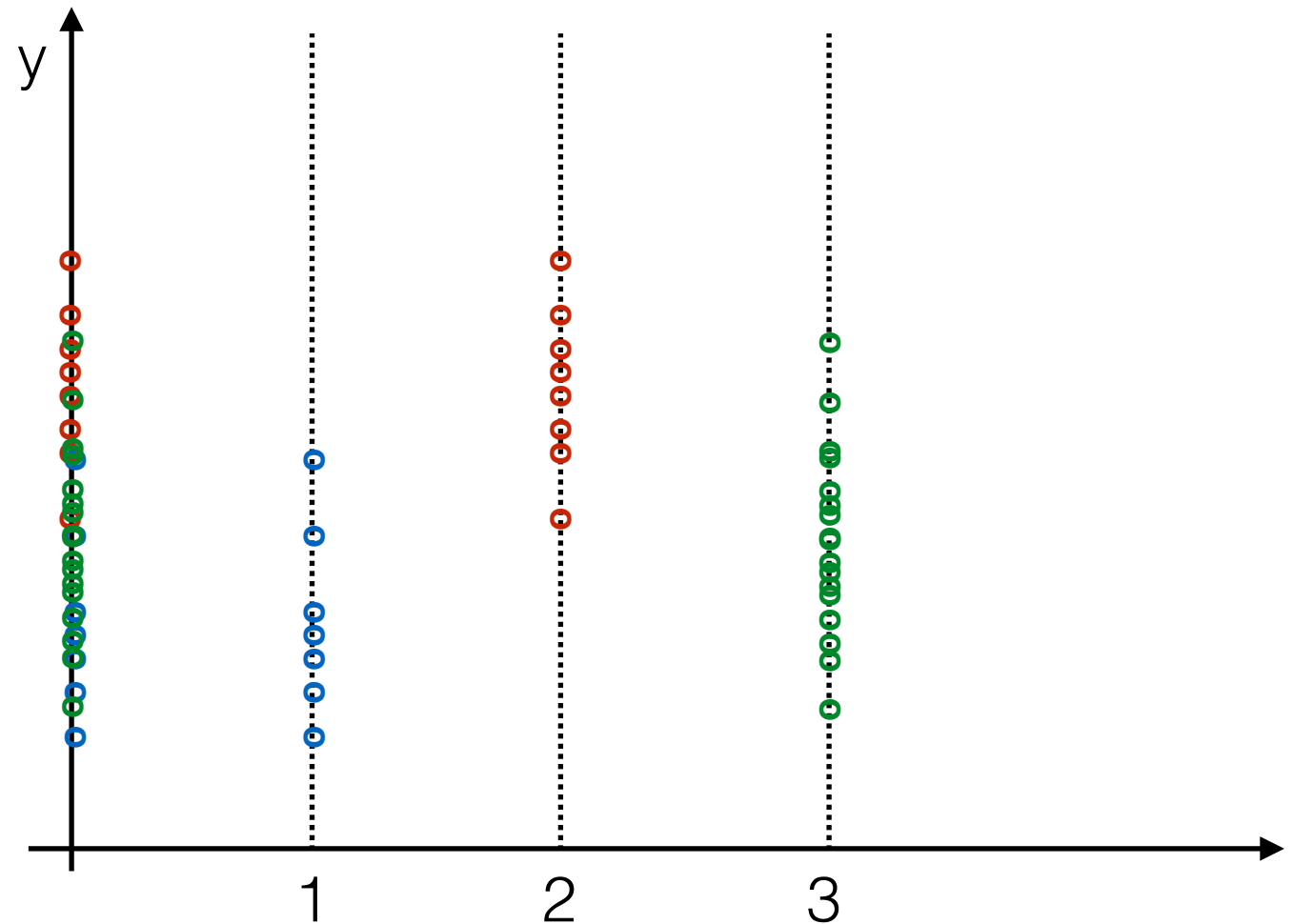
2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovně jednoho faktoru)

k úrovní = k skupin

v každé skupině n měření



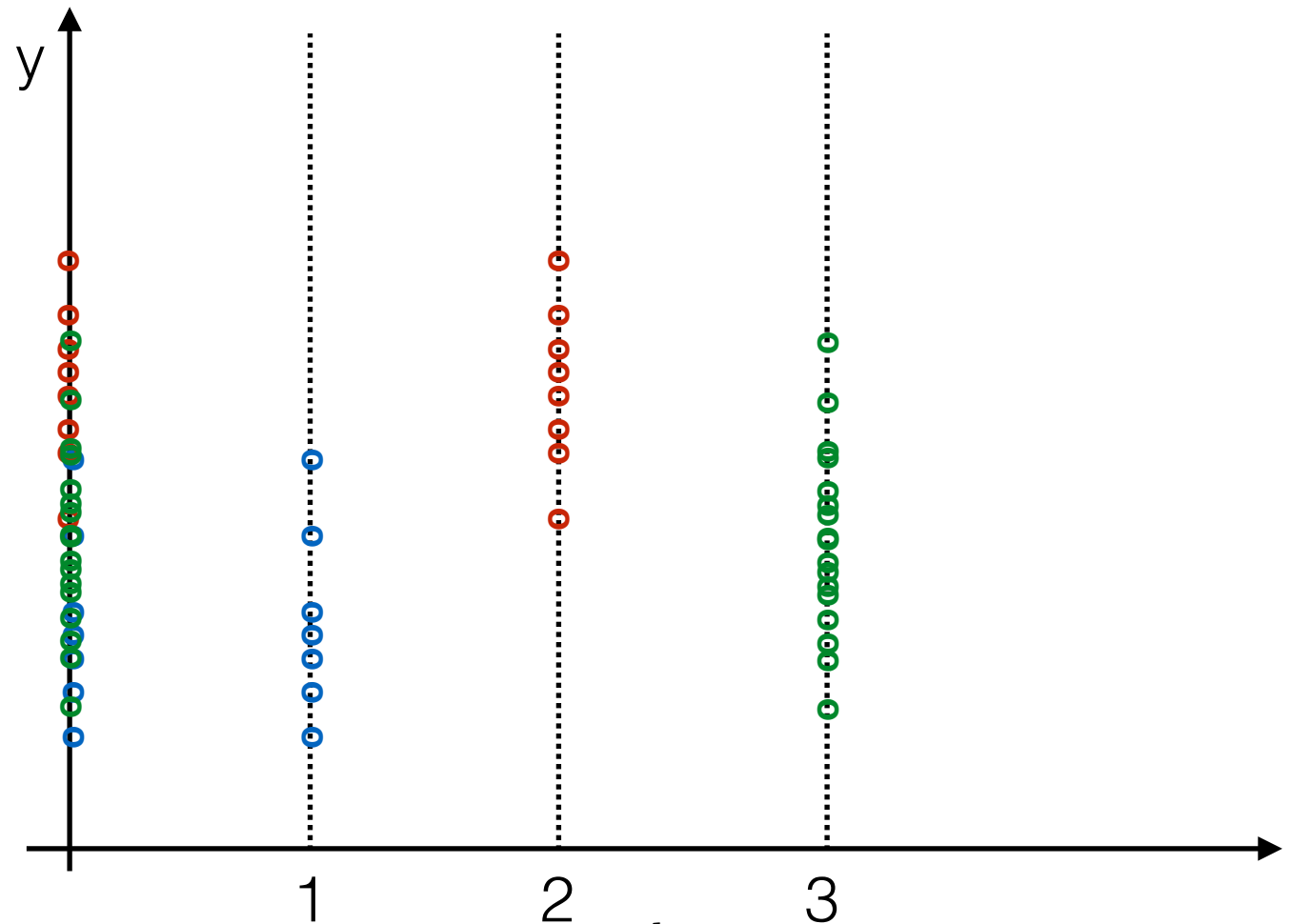
2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovně jednoho faktoru)

k úrovně = k skupin

v každé skupině n měření



1) celkový průměr všech naměřených hodnot:

$$\bar{y} = \frac{1}{n \cdot k} \sum_{i=1}^k \sum_{j=1}^n y_{ij}$$

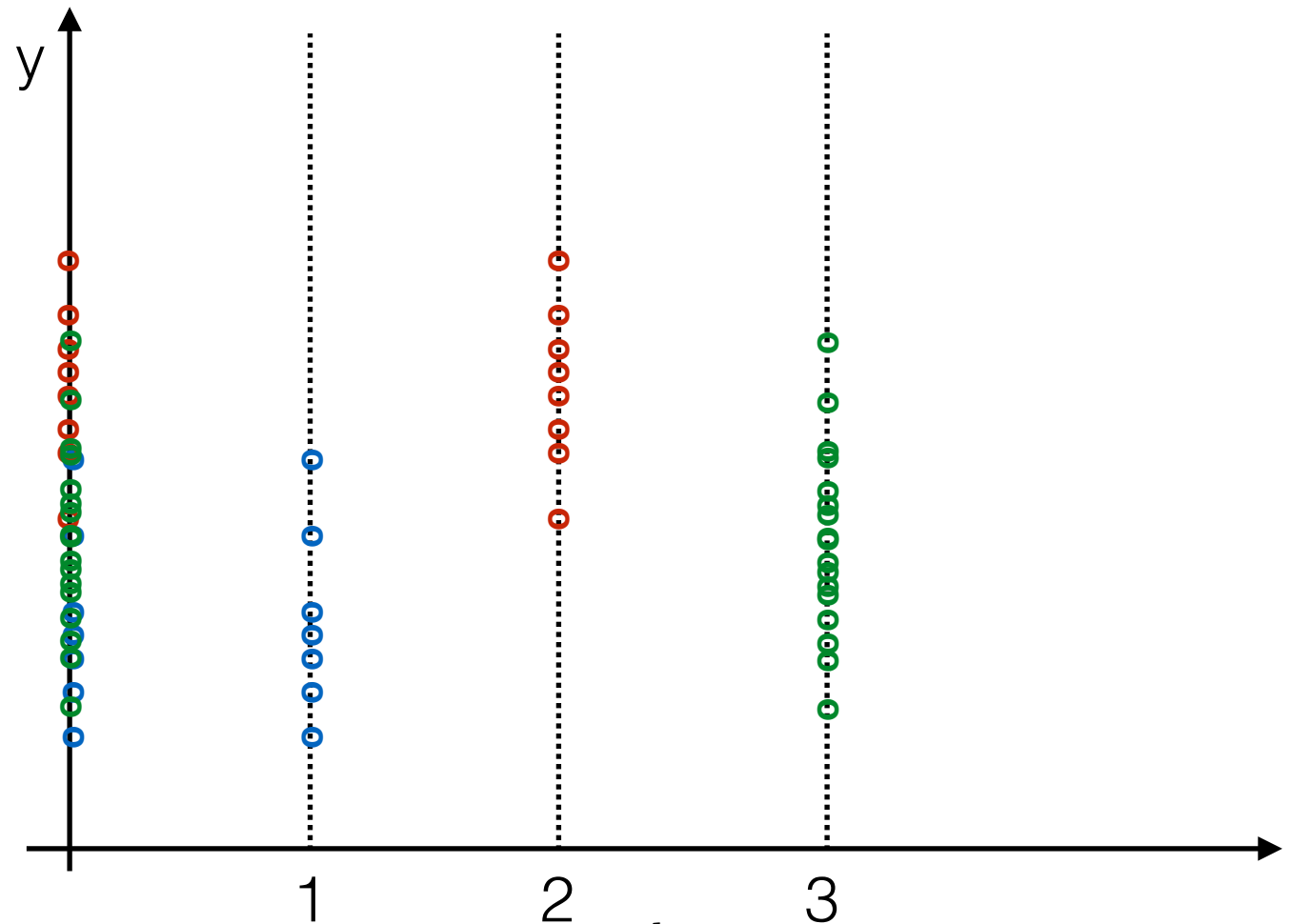
2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovní jednoho faktoru)

k úrovní = k skupin

v každé skupině n měření



1) celkový průměr všech naměřených hodnot: $\bar{y} = \frac{1}{n \cdot k} \sum_{i=1}^k \sum_{j=1}^n y_{ij}$

2) průměry v jednotlivých skupinách: $\bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}$

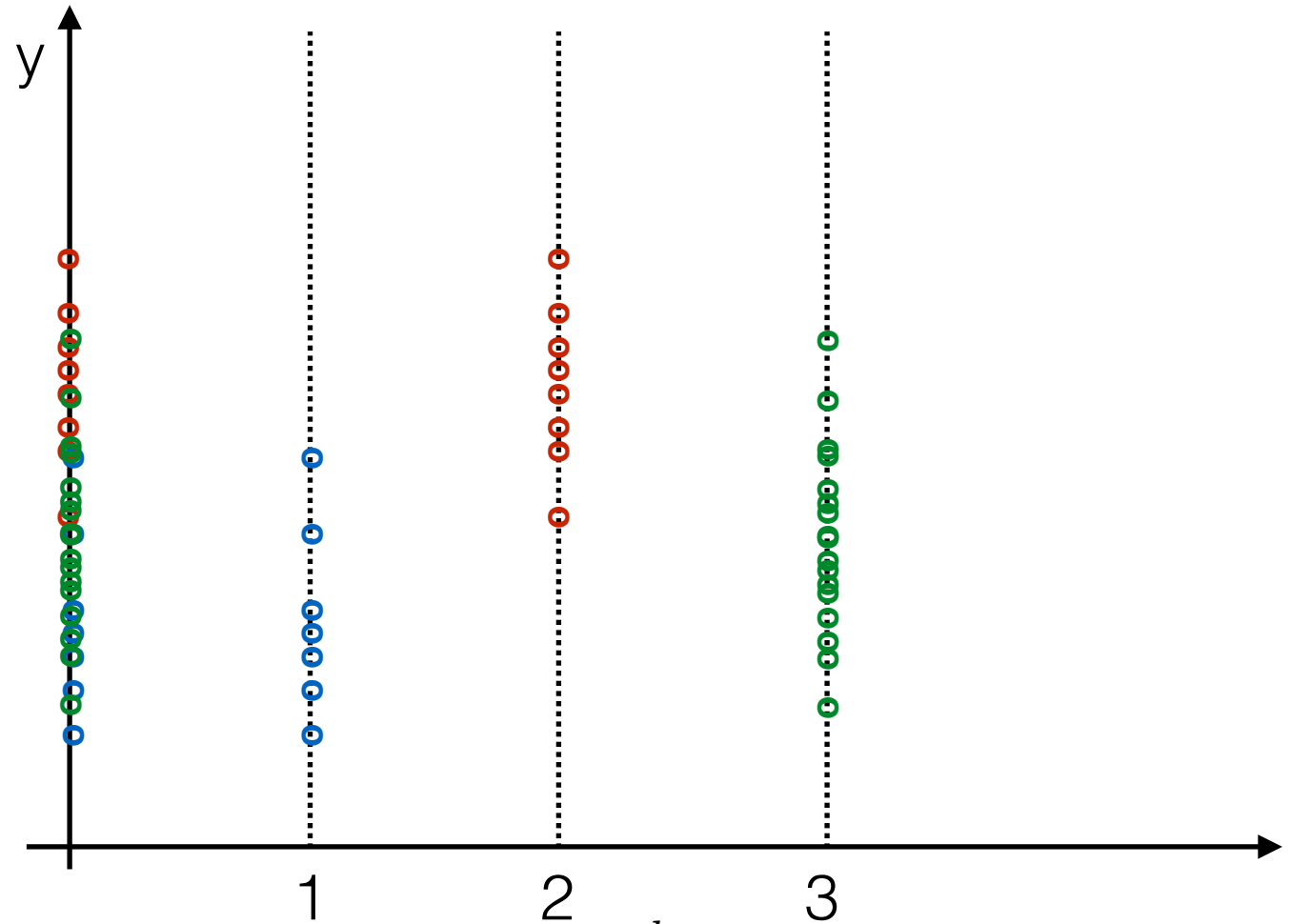
2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovně jednoho faktoru)

k úrovně = k skupin

v každé skupině n měření



1) celkový průměr všech naměřených hodnot: $\bar{y} = \frac{1}{n \cdot k} \sum_{i=1}^k \sum_{j=1}^n y_{ij}$

2) průměry v jednotlivých skupinách: $\bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}$

3) celkový součet čtvercových odchylek: $SS_T = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2$

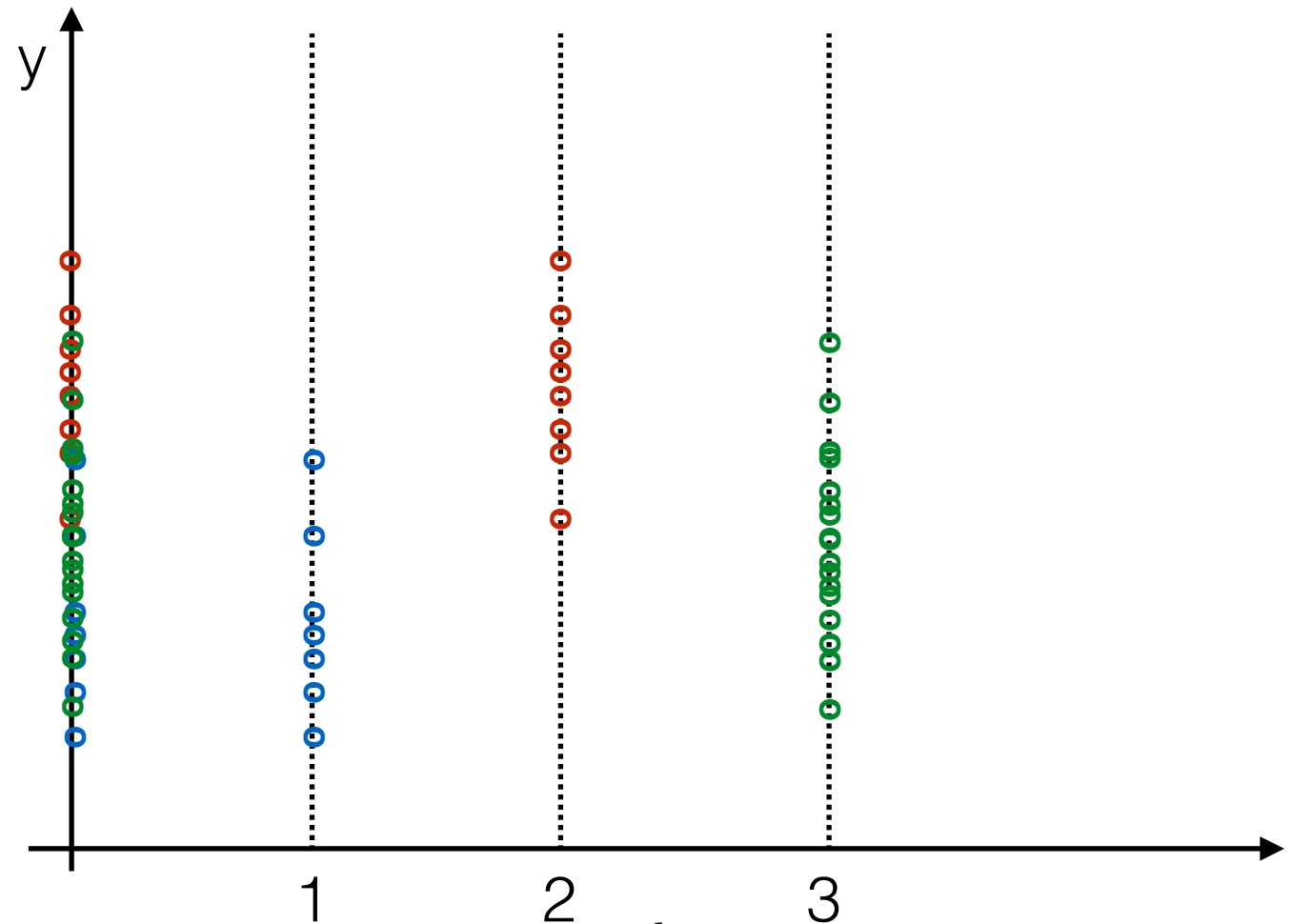
2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovní jednoho faktoru)

k úrovní = k skupin

v každé skupině n měření



1) celkový průměr všech naměřených hodnot: $\bar{y} = \frac{1}{n \cdot k} \sum_{i=1}^k \sum_{j=1}^n y_{ij}$

2) průměry v jednotlivých skupinách: $\bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}$

3) celkový součet čtvercových odchylek: $SS_T = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2$

4) součet čtvercových odchylek uvnitř skupin $SS_E = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$

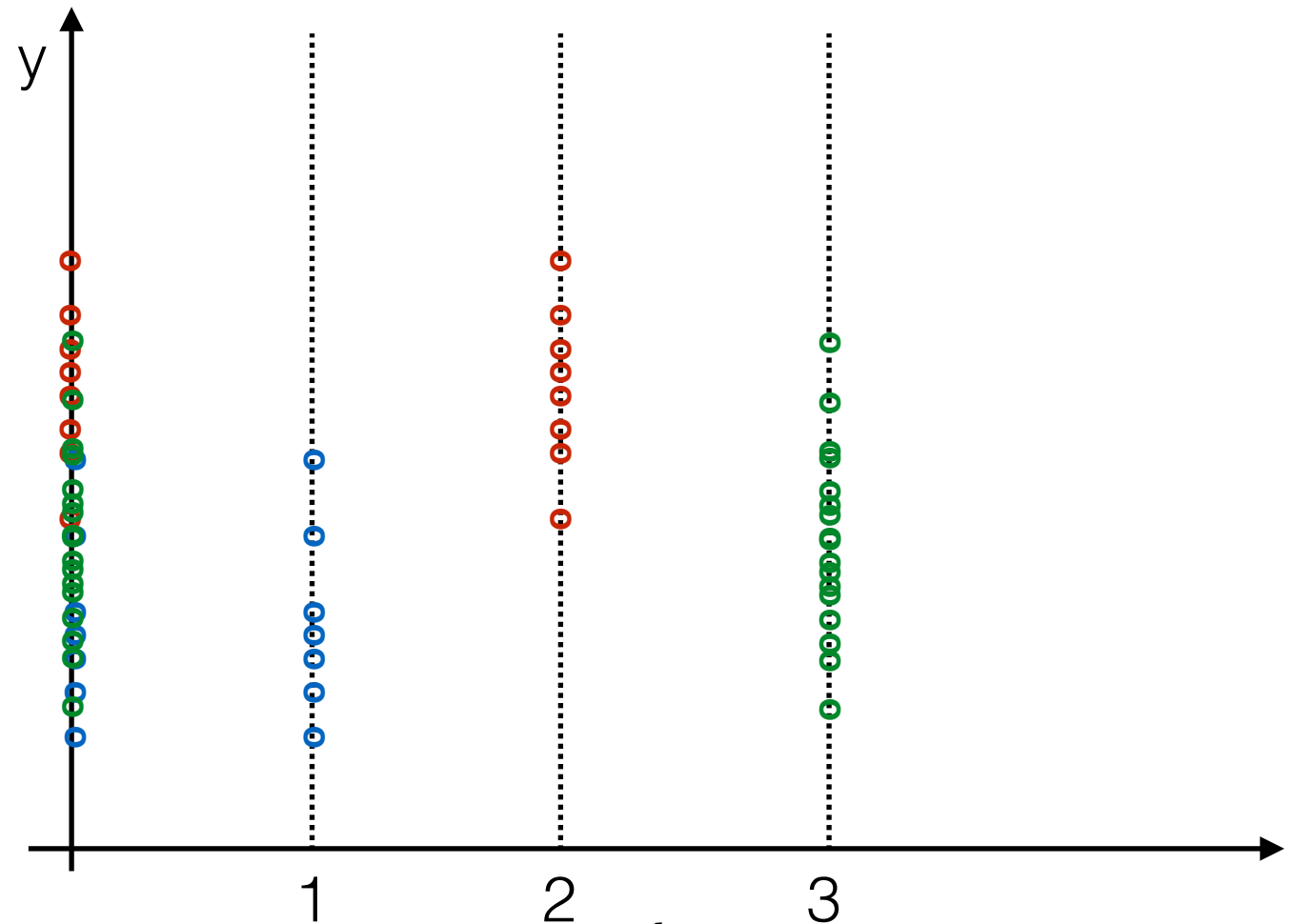
2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

pro 1 faktor = všechna data rozdělíme do skupin podle jednoho hlediska (úrovně jednoho faktoru)

k úrovní = k skupin

v každé skupině n měření



1) celkový průměr všech naměřených hodnot: $\bar{y} = \frac{1}{n \cdot k} \sum_{i=1}^k \sum_{j=1}^n y_{ij}$

2) průměry v jednotlivých skupinách: $\bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}$

3) celkový součet čtvercových odchylek: $SS_T = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2$

4) součet čtvercových odchylek uvnitř skupin $SS_E = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$

5) součet čtvercových odchylek mezi skupinami $SS_F = n \sum_{i=1}^k (\bar{y}_i - \bar{y})^2$

2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Testová statistika	p-hodnota
----------------------	-------------------	--------------------	---------------------	-----------------------	-----------

2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Testová statistika	p-hodnota
faktor (mezi skupinami)	SS_F	$k-1$	$MS_F = \frac{SS_F}{k-1}$		

2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Testová statistika	p-hodnota
faktor (mezi skupinami)	SS_F	$k-1$	$MS_F = \frac{SS_F}{k-1}$		
rezidua (uvnitř skupin)	SS_E	$k(n-1)$	$MS_E = \frac{SS_E}{k(n-1)}$		

2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Testová statistika	p-hodnota
faktor (mezi skupinami)	SS_F	$k-1$	$MS_F = \frac{SS_F}{k-1}$		
rezidua (uvnitř skupin)	SS_E	$k(n-1)$	$MS_E = \frac{SS_E}{k(n-1)}$		
celkový	SS_T	$kn-1$			

2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Testová statistika	p-hodnota
faktor (mezi skupinami)	SS_F	$k-1$	$MS_F = \frac{SS_F}{k-1}$	$F = \frac{MS_F}{MS_E}$	
rezidua (uvnitř skupin)	SS_E	$k(n-1)$	$MS_E = \frac{SS_E}{k(n-1)}$		
celkový	SS_T	$kn-1$			

2.3.1. Analýza rozptylu

ANalysis Of VAriance = ANOVA

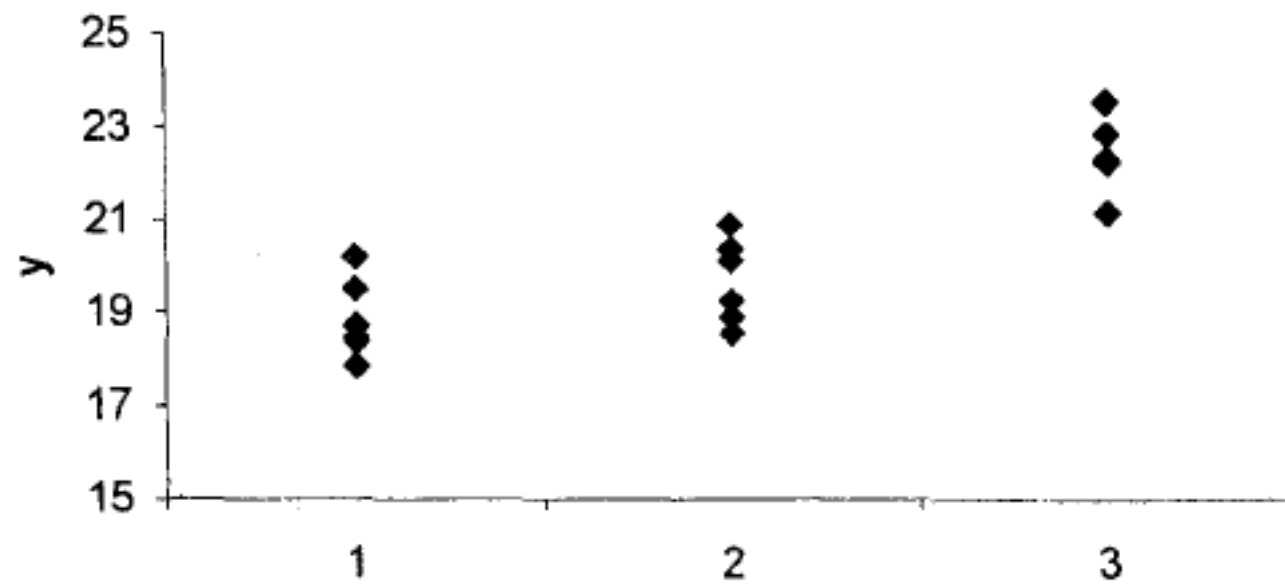
Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Testová statistika	p-hodnota
faktor (mezi skupinami)	SS_F	$k-1$	$MS_F = \frac{SS_F}{k-1}$	$F = \frac{MS_F}{MS_E}$	
rezidua (uvnitř skupin)	SS_E	$k(n-1)$	$MS_E = \frac{SS_E}{k(n-1)}$		
celkový	SS_T	$kn-1$			

testová statistika $F = \frac{\frac{SS_F}{k-1}}{\frac{SS_E}{k(n-1)}} = \frac{MS_F}{MS_E}$ má Fisherovo-Snedecorovo rozdělení o $(k-1)$ a $k(n-1)$ stupních volnosti

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

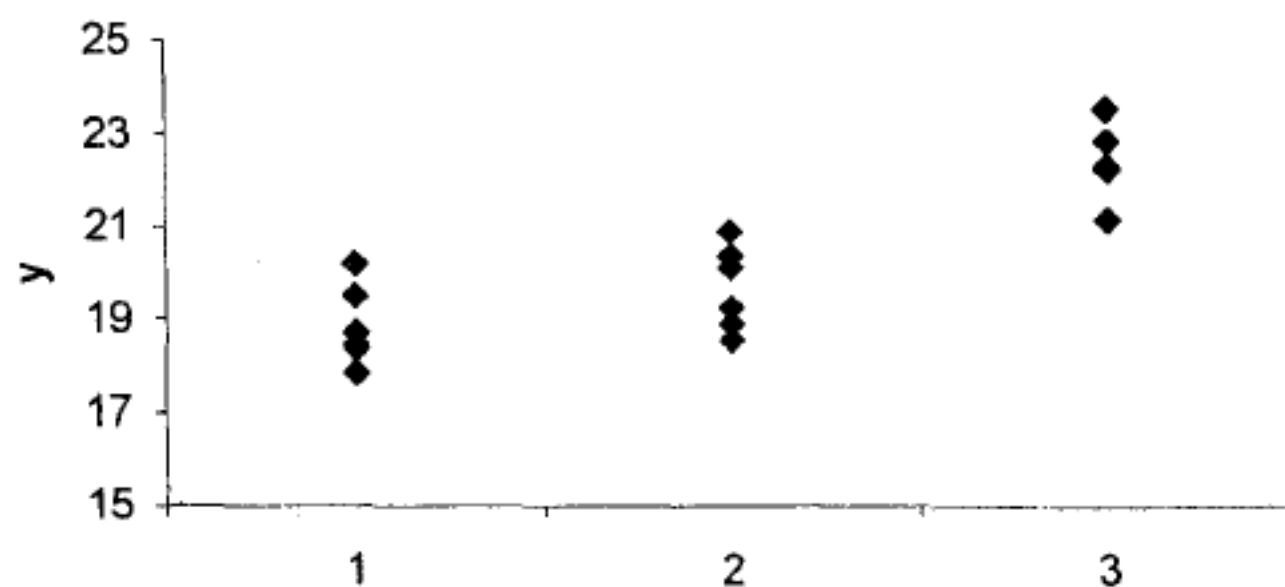
Metoda vyhodnocení: ANOVA pro 1 faktor



2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

Metoda vyhodnocení: ANOVA pro 1 faktor



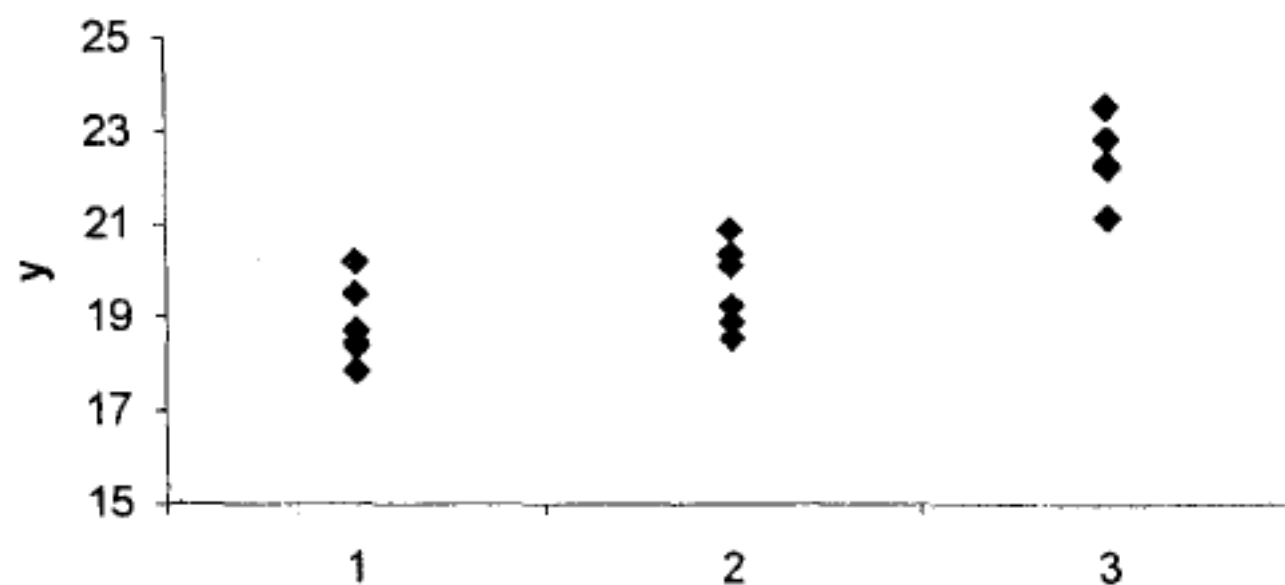
ANOVA

Zdroj variability	SS	St. vol.	MS	F	Hodnota P	F krit
faktor A	40,52333	2	20,26167	29,0513	6,94E-06	3,682317
reziduální	10,46167	15	0,697444			
celkový	50,985	17				

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

Metoda vyhodnocení: ANOVA pro 1 faktor



ANOVA

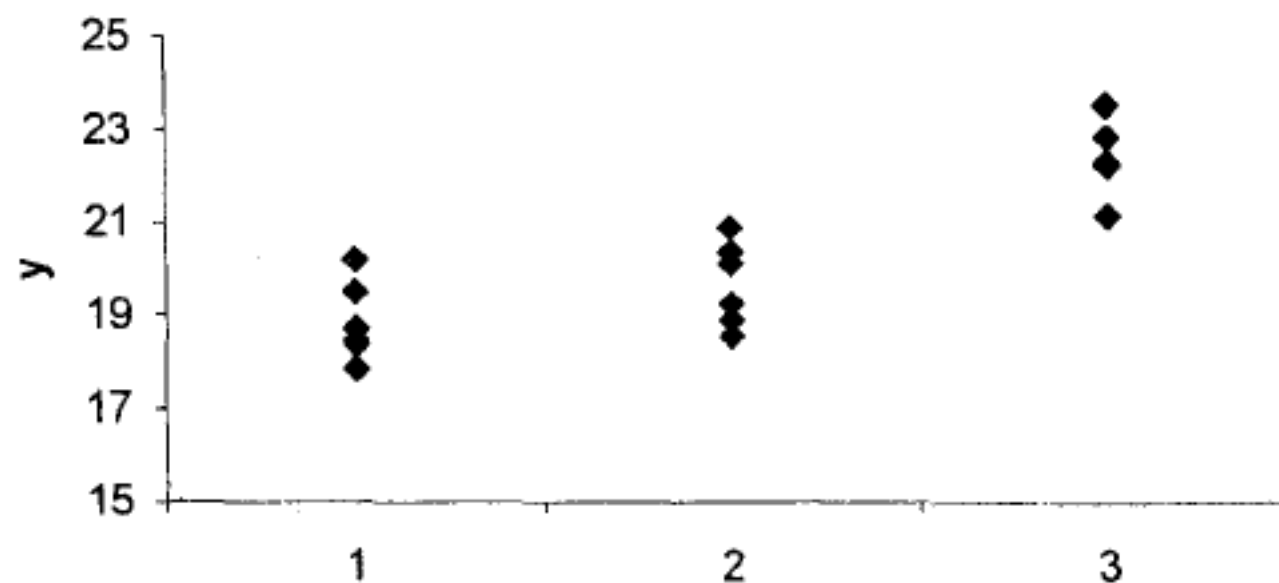
Zdroj variability	SS	St. vol.	MS	F	Hodnota P	F krit
faktor A	40,52333	2	20,26167	29,0513	6,94E-06	3,682317
reziduální	10,46167	15	0,697444			
celkový	50,985	17				

Závěr: nulovou hypotézu zamítáme. Pevnost vláken různých dodavatelů se statisticky významně liší.

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

Metoda vyhodnocení: ANOVA pro 1 faktor



ANOVA

Zdroj variability	SS	St. vol.	MS	F	Hodnota P	F krit
faktor A	40,52333	2	20,26167	29,0513	6,94E-06	3,682317
reziduální	10,46167	15	0,697444			
celkový	50,985	17				

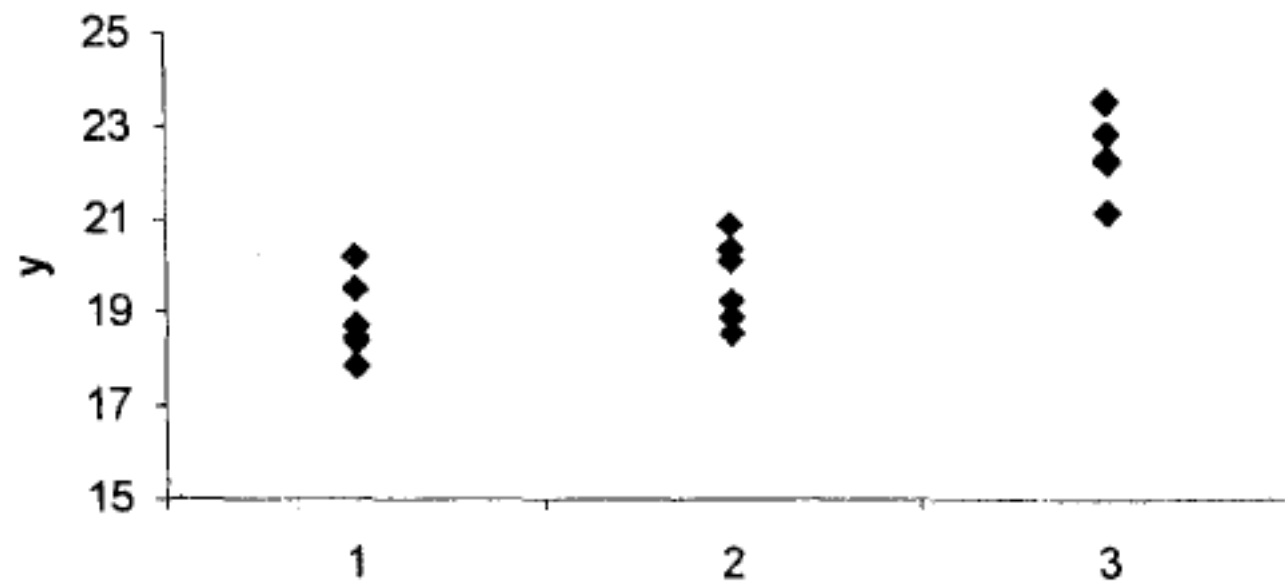
Závěr: nulovou hypotézu zamítáme. Pevnost vláken různých dodavatelů se statisticky významně liší.

Jak se liší dodavatelé mezi sebou?

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

Metoda vyhodnocení: ANOVA pro 1 faktor

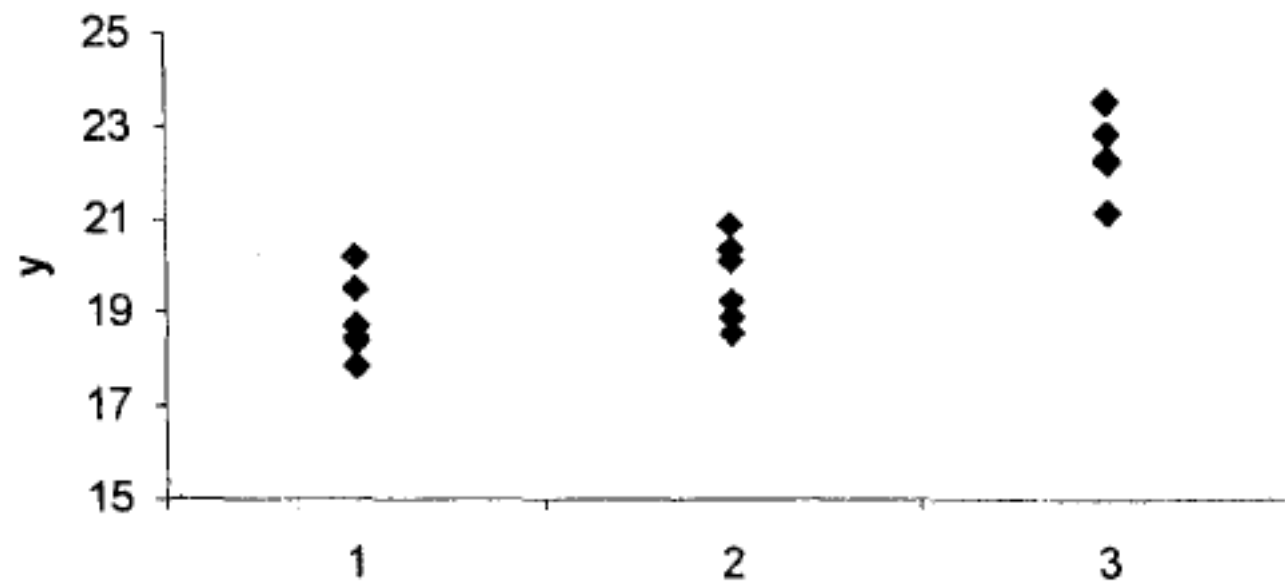


Bonferroniho metoda mnohonásobného srovnání úrovní faktorů:

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

Metoda vyhodnocení: ANOVA pro 1 faktor

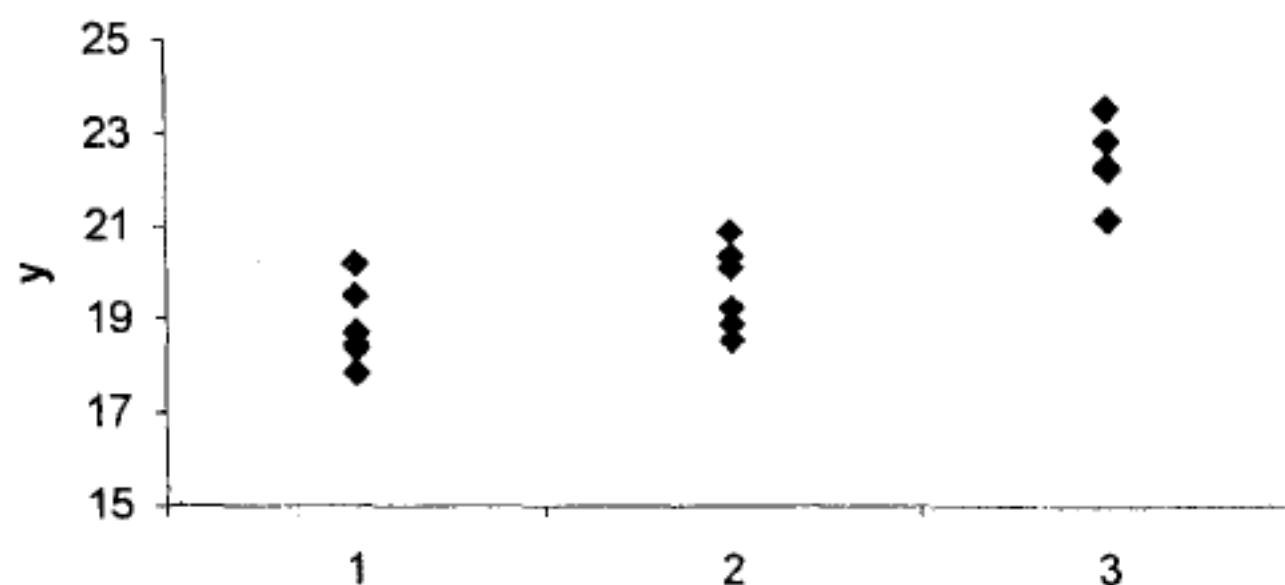


Bonferroniho metoda mnohonásobného srovnání úrovní faktorů: $\bar{y}_i \pm t_{1-\alpha/2p} \cdot \sqrt{\frac{MS_E}{2 \cdot a}}$

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

Metoda vyhodnocení: ANOVA pro 1 faktor



Bonferroniho metoda mnohonásobného srovnání úrovní faktorů:

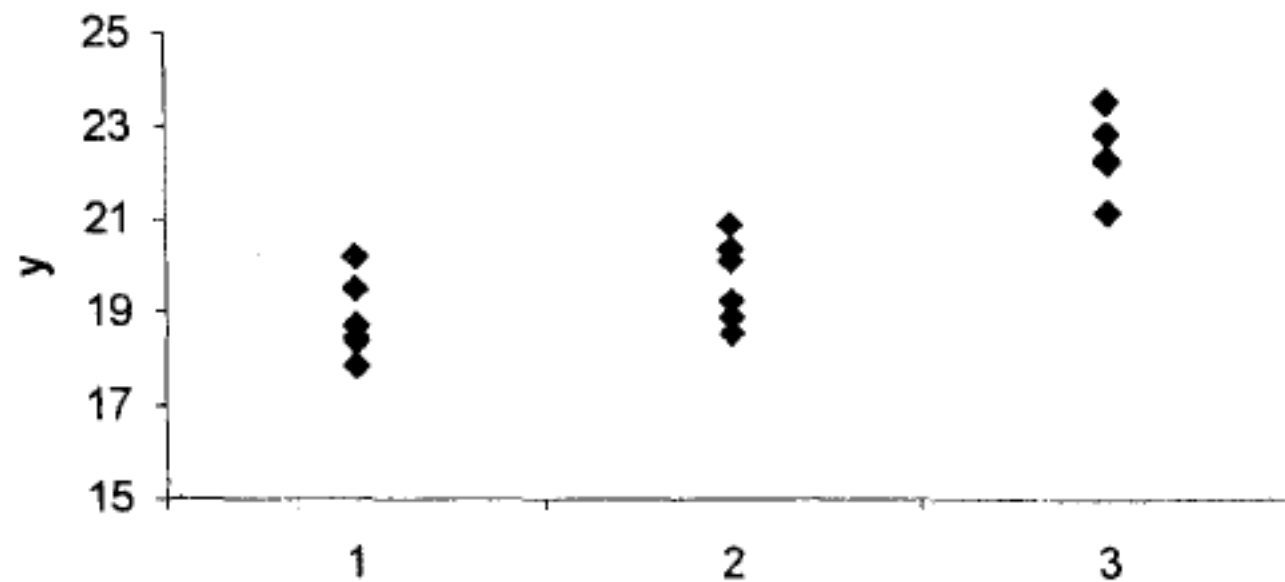
$$\bar{y}_i \pm t_{1-\alpha/2p} \cdot \sqrt{\frac{MS_E}{2 \cdot a}}$$

Úroveň	Průměr	Dolní mez	Horní mez
A ₁	18,867	18,217	19,516
A ₂	19,700	19,051	20,349
A ₃	22,383	21,734	23,033

2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

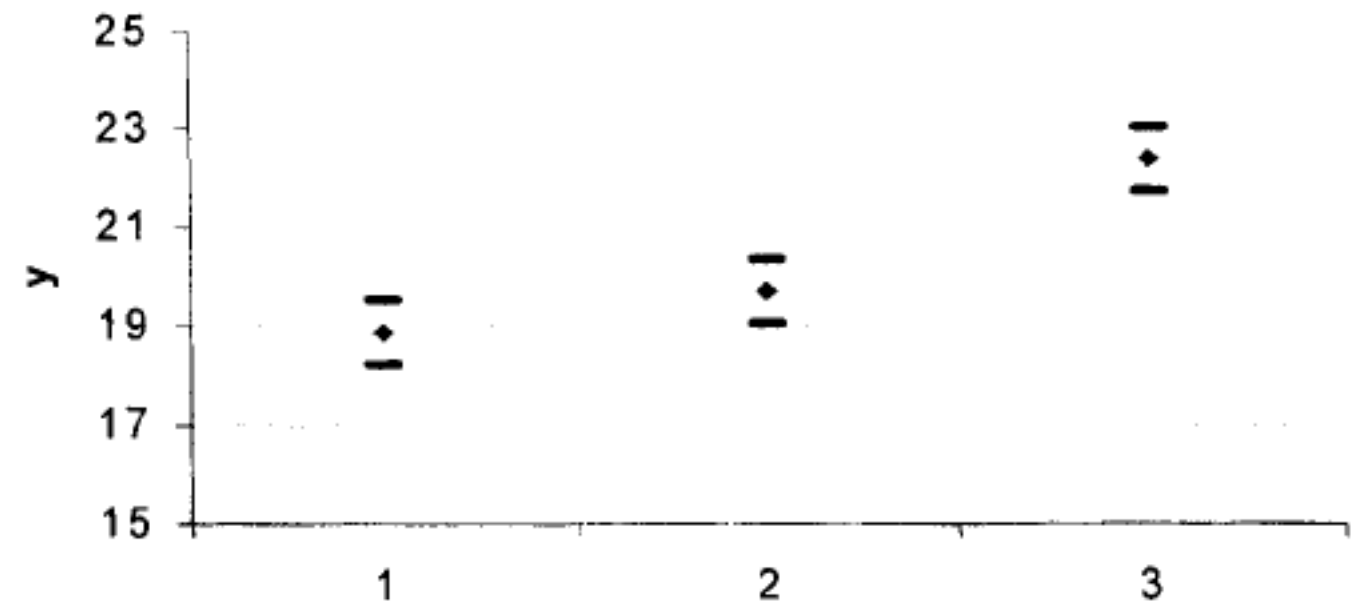
Metoda vyhodnocení: ANOVA pro 1 faktor



Bonferroniho metoda mnohonásobného srovnání úrovní faktorů:

$$\bar{y}_i \pm t_{1-\alpha/2p} \cdot \sqrt{\frac{MS_E}{2 \cdot a}}$$

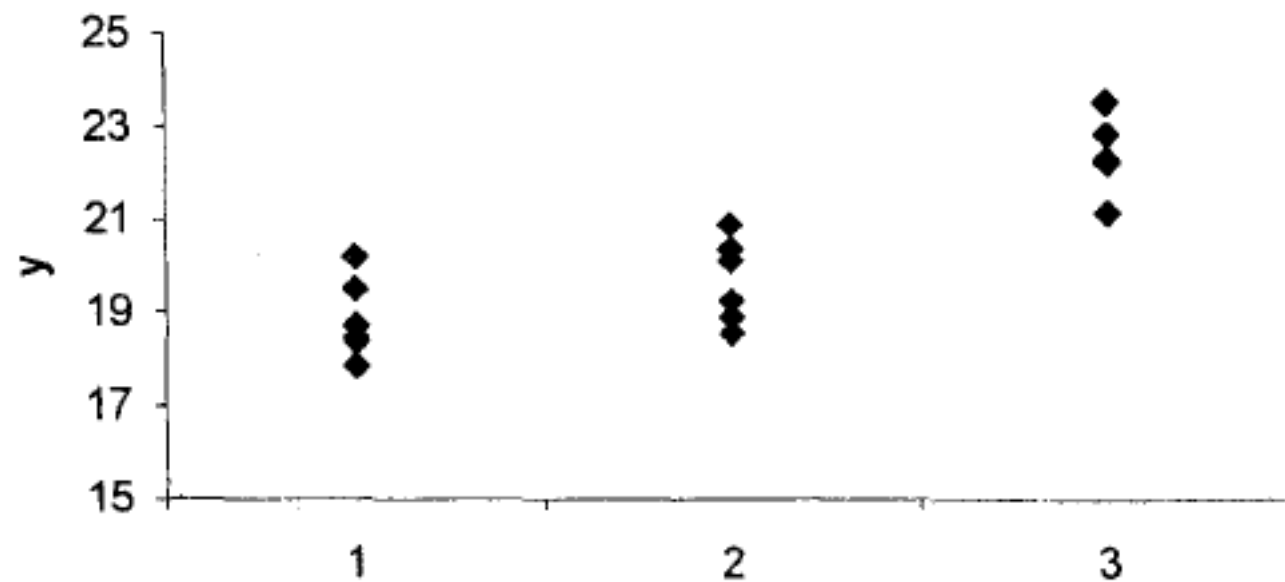
Úroveň	Průměr	Dolní mez	Horní mez
A ₁	18,867	18,217	19,516
A ₂	19,700	19,051	20,349
A ₃	22,383	21,734	23,033



2.3.1. Analýza rozptylu

Srovnání pevnosti vláken od tří dodavatelů

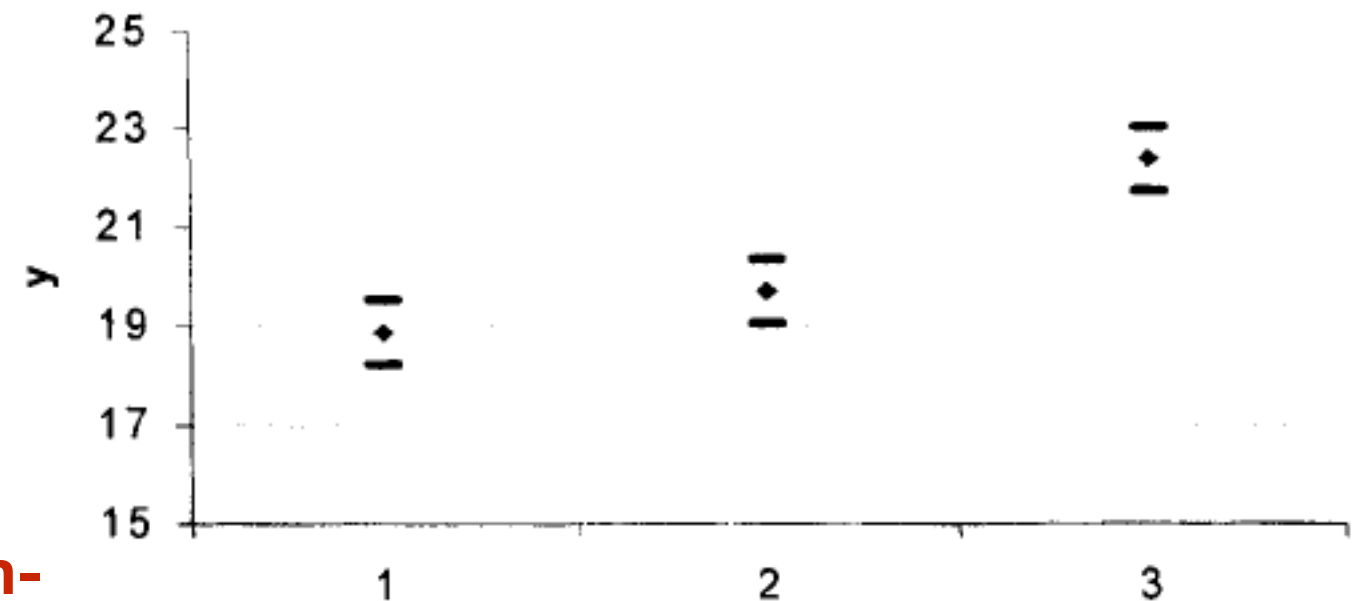
Metoda vyhodnocení: ANOVA pro 1 faktor



Bonferroniho metoda mnohonásobného srovnání úrovní faktorů:

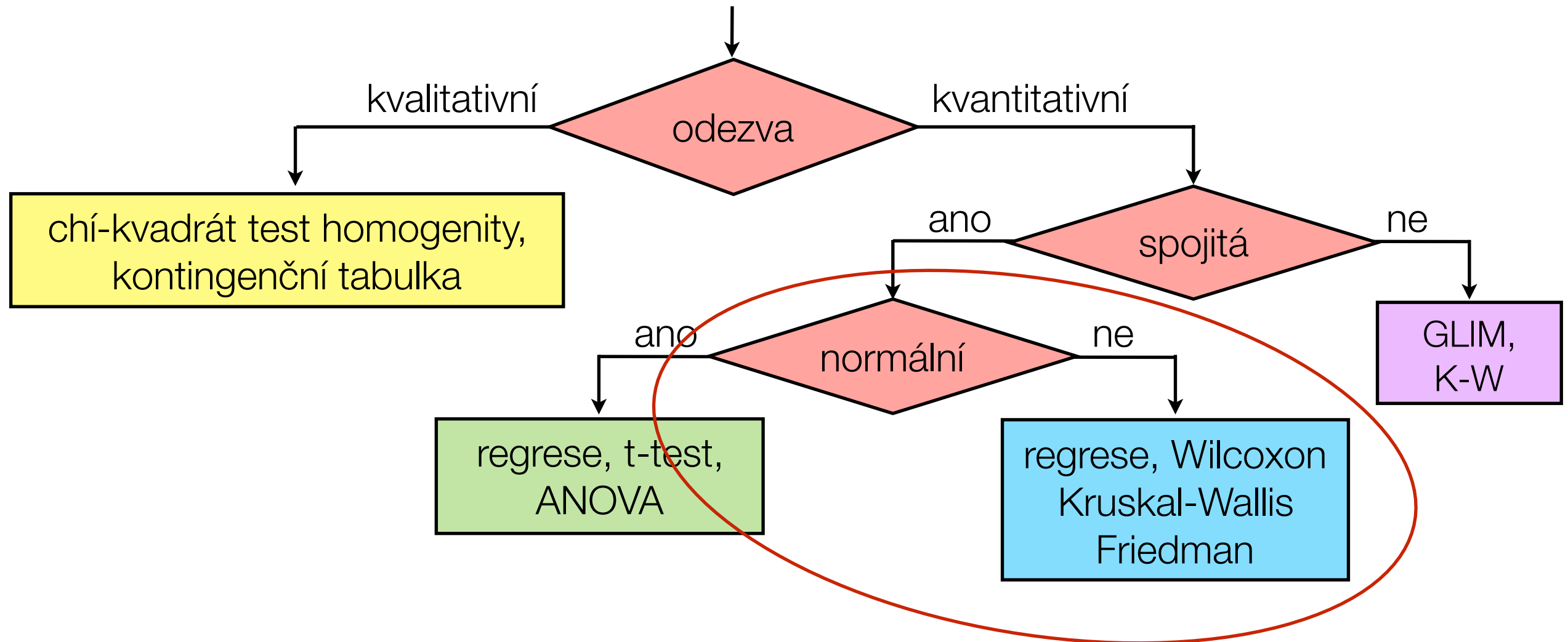
$$\bar{y}_i \pm t_{1-\alpha/2p} \cdot \sqrt{\frac{MS_E}{2 \cdot a}}$$

Úroveň	Průměr	Dolní mez	Horní mez
A ₁	18,867	18,217	19,516
A ₂	19,700	19,051	20,349
A ₃	22,383	21,734	23,033

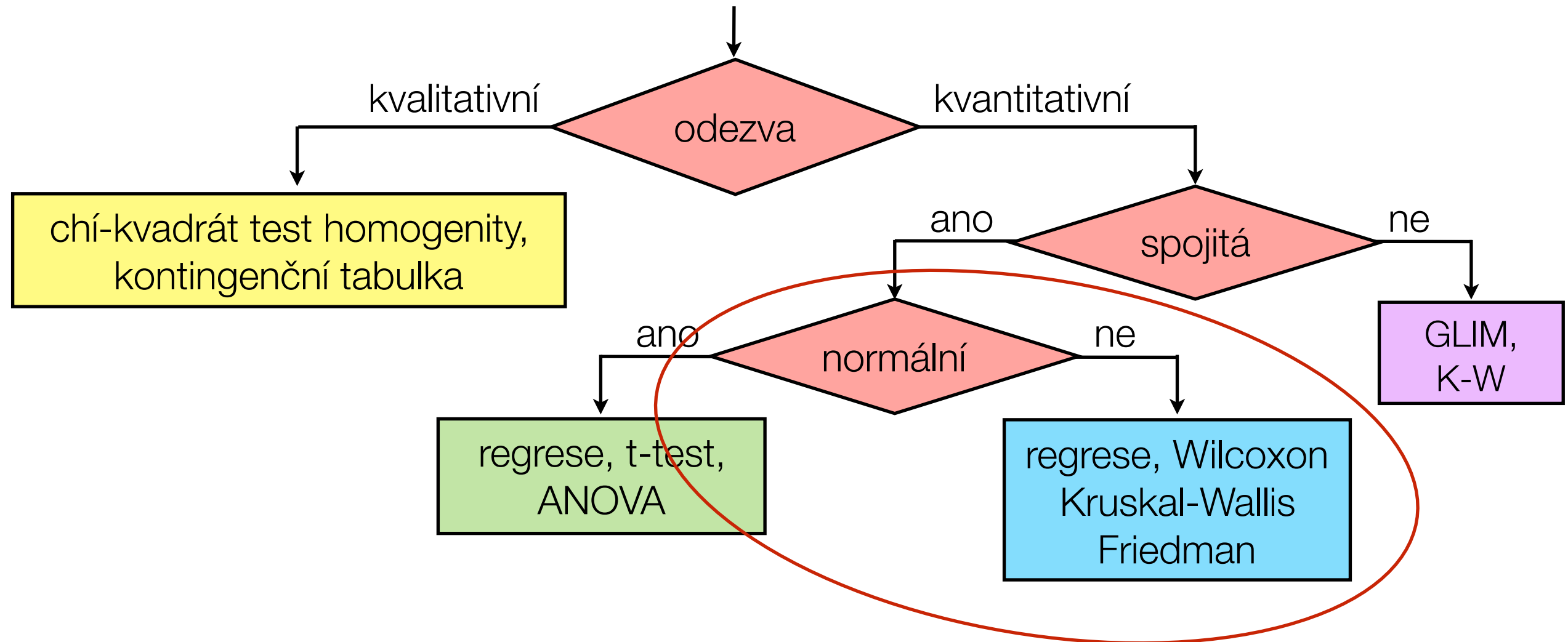


=> dodavatele 1 a 2 nelze statisticky odlišit, dodavatel 3 je významně lepší (jeho vlákna dosahují statisticky významně vyšší pevnosti než od dodavatelů 1 a 2).

2.3. Kvantitativní odezva, jeden faktor, více úrovní



2.3. Kvantitativní odezva, jeden faktor, více úrovní



2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Kruskal – Wallisův test

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly pořízeny ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly pořízeny ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.
- Nulovou hypotézu H_0 můžeme zapsat takto
$$H_0: F_1(x) = F_2(x) = \dots = F_k(x) \text{ pro všechna } x.$$

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly pořízeny ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.
- Nulovou hypotézu H_0 můžeme zapsat takto
$$H_0: F_1(x) = F_2(x) = \dots = F_k(x) \text{ pro všechna } x.$$

Postup při stanovení testového kritéria:

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly pořízeny ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.
- Nulovou hypotézu H_0 můžeme zapsat takto
$$H_0: F_1(x) = F_2(x) = \dots = F_k(x) \text{ pro všechna } x.$$

Postup při stanovení testového kritéria:

- 1) Máme k dispozici k výběrových souborů o četnostech $n_1, n_2, \dots, n_k$.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly poríženy ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.
- Nulovou hypotézu H_0 můžeme zapsat takto
$$H_0: F_1(x) = F_2(x) = \dots = F_k(x) \text{ pro všechna } x.$$

Postup při stanovení testového kritéria:

- 1) Máme k dispozici k výběrových souborů o četnostech $n_1, n_2, \dots, n_k$.
- 2) Všechny výběrové soubory sloučíme do jediného souboru.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly pořízeny ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.
- Nulovou hypotézu H_0 můžeme zapsat takto
$$H_0: F_1(x) = F_2(x) = \dots = F_k(x) \text{ pro všechna } x.$$

Postup při stanovení testového kritéria:

- 1) Máme k dispozici k výběrových souborů o četnostech $n_1, n_2, \dots, n_k$.
- 2) Všechny výběrové soubory sloučíme do jediného souboru.
- 3) Každé hodnotě souboru přiřadíme vzestupně pořadové číslo, stejným hodnotám pak pořadí průměrné.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Kruskal – Wallisův test

- Je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.
- Slouží k ověření nulové hypotézy H_0 , že $k > 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \dots, n_k$ pochází z jednoho základního souboru.
- Předpokládáme, že tyto náhodné výběry byly pořízeny ze základních souborů se spojitými distribučními funkcemi $F_1(x), F_2(x), \dots, F_k(x)$.
- Nulovou hypotézu H_0 můžeme zapsat takto
$$H_0: F_1(x) = F_2(x) = \dots = F_k(x) \text{ pro všechna } x.$$

Postup při stanovení testového kritéria:

- 1) Máme k dispozici k výběrových souborů o četnostech $n_1, n_2, \dots, n_k$.
- 2) Všechny výběrové soubory sloučíme do jediného souboru.
- 3) Každé hodnotě souboru přiřadíme vzestupně pořadové číslo, stejným hodnotám pak pořadí průměrné.
- 4) Následně sečteme pořadová čísla jednotlivých pozorování pro každý původní výběrový soubor zvlášť a získáme součty $T_1, T_2, \dots, T_k$
(T_i ; $i = 1, \dots, k$, je tedy součet pořadových čísel pro i -tý výběr)

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

Testová statistika má tvar:
$$KW = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{T_i^2}{n_i} - 3(n+1)$$
 kde $n = n_1 + n_2 + \dots + n_k$.

Statistika KW má za platnosti H_0 při $n_i \rightarrow \infty$ asymptoticky χ^2 – rozdělení o $k-1$ stupních volnosti.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

Testová statistika má tvar:
$$KW = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{T_i^2}{n_i} - 3(n+1)$$
 kde $n = n_1 + n_2 + \dots + n_k$.

Statistika KW má za platnosti H_0 při $n_i \rightarrow \infty$ asymptoticky χ^2 – rozdělení o $k-1$ stupních volnosti.

Pokud $KW > \chi_{\alpha}^2(n-1)$, přijímáme hypotézu alternativní, podle které se hodnoty nejméně dvou porovnávaných výběrových souborů od sebe statisticky významně liší.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením $\Rightarrow$ nelze použít ANOVA.

Kruskal – Wallisův test

Testová statistika má tvar: $KW = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{T_i^2}{n_i} - 3(n+1)$
kde $n = n_1 + n_2 + \dots + n_k$.

Statistika KW má za platnosti H_0 při $n_i \rightarrow \infty$ asymptoticky χ^2 – rozdělení o $k-1$ stupních volnosti.

Pokud $KW > \chi_{\alpha}^2(n-1)$, přijímáme hypotézu alternativní, podle které se hodnoty nejméně dvou porovnávaných výběrových souborů od sebe statisticky významně liší.

Jestliže se v posloupnosti zjištěných údajů vyskytnou shodné hodnoty, kterým se přiřazuje průměrné pořadí, je nutno hodnotu KW dělit korekčním faktorem

$$K = 1 - \frac{1}{n^3 - n} \sum_{j=1}^p (t_j^3 - t_j)$$

kde p je počet tříd se stejným pořadím a t_i počet pořadí v i -té třídě.

Opravené testové kritérium se stanoví jako $KW_{opr} = \frac{KW}{K}$

2.3.2. Neparametrická analýza rozptylu

The screenshot shows the Minitab software interface. The 'Stat' menu is open, and 'Nonparametrics' is selected. The 'Session' window displays the results of a Kruskal-Wallis Test on C1 versus C3.

Stat Menu Options:

- Basic Statistics
- Regression
- ANOVA
- DOF
- Control Charts
- Quality Tools
- Reliability/Survival
- Multivariate
- Time Series
- Tables
- Nonparametrics**
- EDA
- Power and Sample Size

Session Window Output:

Kruskal-Wallis Test

C3	N	Median
1	2	16,95
2	2	16,68
3	2	17,18
4	2	16,47
5	2	16,55
Overall	10	

H = 6,44 DF = 4

* NCTE * One or more small samples

Kruskal-Wallis Test: C1 versus C3

Kruskal-Wallis Test on C1

C3	N	Median	Rank	Ave
1	2	16,95	7,0	0,78
2	2	16,68	4,5	-0,52
3	2	17,18	9,5	2,09
4	2	16,47	3,0	-1,31
5	2	16,55	3,5	-1,04
Overall	10		5,5	

H = 6,44 DF = 4 P = 0,169

* NCTE * One or more small samples

1

2.3.2. Neparametrická analýza rozptylu

The screenshot displays the Minitab software interface. The **Stat** menu is open, showing various statistical tests. The **Nonparametrics** submenu is also open, highlighting the **Kruskal-Wallis...** option. The **Session** window on the left shows the results of a Kruskal-Wallis Test for C1 versus C3. The **Kruskal-Wallis** dialog box is open on the right, with **C1** selected as the Response and **C3** as the Factor. The **Select** button is highlighted.

Stat menu options:

- Basic Statistics
- Regression
- ANOVA
- DOF
- Control Charts
- Quality Tools
- Reliability/Survival
- Multivariate
- Time Series
- Tables
- Nonparametrics**
- EDA
- Power and Sample Size

Nonparametrics submenu options:

- 1-Sample Sign...
- 1-Sample Wilcoxon...
- Mann-Whitney...
- Kruskal-Wallis...**
- Mood's Median Test...
- Friedman...
- Runs Test...
- Pairwise Averages ..
- Pairwise Differences...
- Pairwise Slopes...

Session window output:

Kruskal-Wallis Test

C3	N	Median
1	2	16,95
2	2	16,68
3	2	17,18
4	2	16,47
5	2	16,55
Overall	10	

H = 6,44 DF = 4

* NCTE * One or more small samples

Kruskal-Wallis Test: C1 versus C3

Kruskal-Wallis Test on C1

C3	N	Median	Rank	Ave
1	2	16,95	7,0	0,78
2	2	16,68	4,5	-0,52
3	2	17,18	9,5	2,09
4	2	16,47	3,0	-1,31
5	2	16,55	3,5	-1,04
Overall	10		5,5	

H = 6,44 DF = 4 P = 0,169

* NCTE * One or more small samples

Kruskal-Wallis dialog box:

Response: **C1**

Factor: **C3**

Select

Help OK Cancel

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Neményiho metoda

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Neményiho metoda

- Předpokládáme tzv. „*vyvážený pokusný plán*“, tzn., že všech k výběrů má stejný rozsah (tedy $n_1 = n_2 = \dots = n_k = N$).

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Neményiho metoda

- Předpokládáme tzv. „*vyvážený pokusný plán*“, tzn., že všech k výběrů má stejný rozsah (tedy $n_1 = n_2 = \dots = n_k = N$).
- Spočítáme $\frac{k(k-1)}{2}$ diferencí $|T_i - T_j|$ které porovnáváme s kritickými hodnotami.

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Neményiho metoda

- Předpokládáme tzv. „vyvážený pokusný plán“, tzn., že všech k výběrů má stejný rozsah (tedy $n_1 = n_2 = \dots = n_k = N$).
- Spočítáme $\frac{k(k-1)}{2}$ diferencí $|T_i - T_j|$ které porovnáváme s kritickými hodnotami.
- Kritické hodnoty D_α se hledají pro hladinu významnosti α , pro k porovnávaných tříd a N opakování v každé třídě ($n_1 = n_2 = \dots = n_k = N$).

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Neményiho metoda

- Předpokládáme tzv. „vyvážený pokusný plán“, tzn., že všech k výběrů má stejný rozsah (tedy $n_1 = n_2 = \dots = n_k = N$).
- Spočítáme $\frac{k(k-1)}{2}$ diferencí $|T_i - T_j|$ které porovnáváme s kritickými hodnotami.
- Kritické hodnoty D_α se hledají pro hladinu významnosti α , pro k porovnávaných tříd a N opakování v každé třídě ($n_1 = n_2 = \dots = n_k = N$).
- Je-li nějaká difference větší nebo rovna kritické hodnotě D_α pro Neményiho metodu, zamítá se hypotéza o neprůkaznosti difference, tzn. že i -tý a j -tý výběr pocházejí z téhož rozdělení.

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Dunnova metoda

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Dunnova metoda

- Použijeme ji v případě „nevyváženého pokusného plánu“, tzn., že výběry nemají stejný rozsah (tedy nemusí být $n_1 = n_2 = \dots = n_k$).

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Dunnova metoda

- Použijeme ji v případě „nevyváženého pokusného plánu“, tzn., že výběry nemají stejný rozsah (tedy nemusí být $n_1 = n_2 = \dots = n_k$).
- Spočítáme $\frac{k(k-1)}{2}$ diferencí $|T_i - T_j|$ které porovnáváme s kritickými hodnotami.

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Dunnova metoda

- Použijeme ji v případě „nevyváženého pokusného plánu“, tzn., že výběry nemají stejný rozsah (tedy nemusí být $n_1 = n_2 = \dots = n_k$).
- Spočítáme $\frac{k(k-1)}{2}$ diferencí $|T_i - T_j|$ které porovnáваме s kritickými hodnotami.
- Pokud pro nějakou diferenci platí

$$|T_i - T_j| > u_\beta \sqrt{\frac{n(n-1)}{12} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

kde u_β je β - kvantil standardního normálního rozdělení, $\beta = \frac{\alpha}{k(k-1)}$, zamítáme na hladině významnosti α hypotézu, že i -tý výběr (s rozsahem n_i) a j -tý výběr (s rozsahem n_j) pocházejí z téhož rozdělení.

2.3.2. Neparametrická analýza rozptylu

Neparametrické metody mnohonásobného porovnávání.

Dunnova metoda

- Použijeme ji v případě „nevyváženého pokusného plánu“, tzn., že výběry nemají stejný rozsah (tedy nemusí být $n_1 = n_2 = \dots = n_k$).
- Spočítáme $\frac{k(k-1)}{2}$ diferencí $|T_i - T_j|$ které porovnáваме s kritickými hodnotami.
- Pokud pro nějakou diferenci platí

$$|T_i - T_j| > u_\beta \sqrt{\frac{n(n-1)}{12} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

kde u_β je β - kvantil standardního normálního rozdělení, $\beta = \frac{\alpha}{k(k-1)}$,
zamítáme na hladině významnosti α hypotézu, že i -tý výběr
(s rozsahem n_i) a j -tý výběr (s rozsahem n_j) pocházejí z téhož rozdělení.

2.3.2. Neparametrická analýza rozptylu

Příklad

Úloha: Ověřit, zda výnosy silážní kukuřice jsou ovlivněny rozdílnou dávkou NPK v hnojení.

2.3.2. Neparametrická analýza rozptylu

Příklad

Úloha: Ověřit, zda výnosy silážní kukuřice jsou ovlivněny rozdílnou dávkou NPK v hnojení.

- Odezva: výnos v t/ha
- Faktor: způsob hnojení (4 úrovně lišící se dávkou NPK)

2.3.2. Neparametrická analýza rozptylu

Příklad

Úloha: Ověřit, zda výnosy silážní kukuřice jsou ovlivněny rozdílnou dávkou NPK v hnojení.

- Odezva: výnos v t/ha
- Faktor: způsob hnojení (4 úrovně lišící se dávkou NPK)

Provedeme polní experiment, v němž budeme sledovat výnosy pro čtyři varianty hnojení silážní kukuřice rozdílnou dávkou NPK v hnojivech, označené jako V_1 až V_4 . Každá varianta bude ověřována na 8 parcelách a budou měřeny výnosy sklizené hmoty v tunách na hektar.

2.3.2. Neparametrická analýza rozptylu

Příklad

Úloha: Ověřit, zda výnosy silážní kukuřice jsou ovlivněny rozdílnou dávkou NPK v hnojení.

- Odezva: výnos v t/ha
- Faktor: způsob hnojení (4 úrovně lišící se dávkou NPK)

Provedeme polní experiment, v němž budeme sledovat výnosy pro čtyři varianty hnojení silážní kukuřice rozdílnou dávkou NPK v hnojivech, označené jako V_1 až V_4 . Každá varianta bude ověřována na 8 parcelách a budou měřeny výnosy sklizené hmoty v tunách na hektar.

H_0 : výnosy jednotlivých variant jsou shodné

H_1 : výnosy jednotlivých variant jsou rozdílné

2.3.2. Neparametrická analýza rozptylu

Příklad

Úloha: Ověřit, zda výnosy silážní kukuřice jsou ovlivněny rozdílnou dávkou NPK v hnojivu.

- Odezva: výnos v t/ha
- Faktor: způsob hnojení (4 úrovně lišící se dávkou NPK)

Provedeme polní experiment, v němž budeme sledovat výnosy pro čtyři varianty hnojení silážní kukuřice rozdílnou dávkou NPK v hnojivech, označené jako V_1 až V_4 . Každá varianta bude ověřována na 8 parcelách a budou měřeny výnosy sklizené hmoty v tunách na hektar.

H_0 : výnosy jednotlivých variant jsou shodné

H_1 : výnosy jednotlivých variant jsou rozdílné

Hypotéza o normalitě napozorovaných dat byla zamítnuta. Proto použijeme Kruskal-Wallisův test.

2.3.2. Neparametrická analýza rozptylu

Příklad

varianta	Výnos kukuřice v t z hektaru na parcele číslo:								součet T_i
	1	2	3	4	5	6	7	8	
V ₁ pořadí	1,29 23,5	1,19 8,5	1,23 15	1,33 28,5	1,27 20	1,29 23,5	1,31 27	1,20 11,5	157,5
V ₂ pořadí	1,30 26	1,33 28,5	1,29 23,5	1,37 31	1,35 30	1,25 18,5	1,38 32	1,29 23,5	213
V ₃ pořadí	1,20 11,5	1,24 16,5	1,25 18,5	1,24 16,5	1,20 11,5	1,21 14	1,28 21	1,17 7	116,5
V ₄ pořadí	1,03 2	1,14 6	1,09 5	1,20 11,5	1,07 4	1,19 8,5	1,01 1	1,05 3	41

2.3.2. Neparametrická analýza rozptylu

Příklad

varianta	Výnos kukuřice v t z hektaru na parcele číslo:								součet T_i
	1	2	3	4	5	6	7	8	
V_1 pořadí	1,29 23,5	1,19 8,5	1,23 15	1,33 28,5	1,27 20	1,29 23,5	1,31 27	1,20 11,5	157,5
V_2 pořadí	1,30 26	1,33 28,5	1,29 23,5	1,37 31	1,35 30	1,25 18,5	1,38 32	1,29 23,5	213
V_3 pořadí	1,20 11,5	1,24 16,5	1,25 18,5	1,24 16,5	1,20 11,5	1,21 14	1,28 21	1,17 7	116,5
V_4 pořadí	1,03 2	1,14 6	1,09 5	1,20 11,5	1,07 4	1,19 8,5	1,01 1	1,05 3	41

$$KW = \frac{12}{32 \cdot 33} \left(\frac{157,5^2}{8} + \frac{213^2}{8} + \frac{116,5^2}{8} + \frac{41^2}{8} \right) - 3 \cdot 33 = 22,3473$$

2.3.2. Neparametrická analýza rozptylu

Příklad

varianta	Výnos kukuřice v t z hektaru na parcele číslo:								součet T_i
	1	2	3	4	5	6	7	8	
V ₁ pořadí	1,29 23,5	1,19 8,5	1,23 15	1,33 28,5	1,27 20	1,29 23,5	1,31 27	1,20 11,5	157,5
V ₂ pořadí	1,30 26	1,33 28,5	1,29 23,5	1,37 31	1,35 30	1,25 18,5	1,38 32	1,29 23,5	213
V ₃ pořadí	1,20 11,5	1,24 16,5	1,25 18,5	1,24 16,5	1,20 11,5	1,21 14	1,28 21	1,17 7	116,5
V ₄ pořadí	1,03 2	1,14 6	1,09 5	1,20 11,5	1,07 4	1,19 8,5	1,01 1	1,05 3	41

$$KW = \frac{12}{32 \cdot 33} \left(\frac{157,5^2}{8} + \frac{213^2}{8} + \frac{116,5^2}{8} + \frac{41^2}{8} \right) - 3 \cdot 33 = 22,3473$$

Vzhledem k výskytu stejných údajů (bylo použito průměrné pořadí) je vhodné opravit testové kritérium korekčním faktorem:

$$K = 1 - \frac{(2^3 - 2) + (4^3 - 4) + (2^3 - 2) + (2^3 - 2) + (4^3 - 4) + (2^3 - 2)}{32^3 - 32} = 0,9956$$

2.3.2. Neparametrická analýza rozptylu

Příklad

varianta	Výnos kukuřice v t z hektaru na parcele číslo:								součet T_i
	1	2	3	4	5	6	7	8	
V ₁ pořadí	1,29 23,5	1,19 8,5	1,23 15	1,33 28,5	1,27 20	1,29 23,5	1,31 27	1,20 11,5	157,5
V ₂ pořadí	1,30 26	1,33 28,5	1,29 23,5	1,37 31	1,35 30	1,25 18,5	1,38 32	1,29 23,5	213
V ₃ pořadí	1,20 11,5	1,24 16,5	1,25 18,5	1,24 16,5	1,20 11,5	1,21 14	1,28 21	1,17 7	116,5
V ₄ pořadí	1,03 2	1,14 6	1,09 5	1,20 11,5	1,07 4	1,19 8,5	1,01 1	1,05 3	41

$$KW = \frac{12}{32 \cdot 33} \left(\frac{157,5^2}{8} + \frac{213^2}{8} + \frac{116,5^2}{8} + \frac{41^2}{8} \right) - 3 \cdot 33 = 22,3473$$

Vzhledem k výskytu stejných údajů (bylo použito průměrné pořadí) je vhodné opravit testové kritérium korekčním faktorem:

$$K = 1 - \frac{(2^3 - 2) + (4^3 - 4) + (2^3 - 2) + (2^3 - 2) + (4^3 - 4) + (2^3 - 2)}{32^3 - 32} = 0,9956$$

$$KW_{opr.} = \frac{22,3473}{0,9956} = 22,446 \quad \chi^2_{0,05(3)} = 7,815 \quad \chi^2_{0,01(3)} = 11,34$$

2.3.2. Neparametrická analýza rozptylu

Příklad

Opravené testové kritérium $KW = 22,446 > \chi^2 \Rightarrow$ na hladině významnosti $\alpha = 0,05$ i $0,01$ přijímáme alternativní hypotézu, podle které se průkazně liší hodnoty výnosu nejméně ve 2 třídách.

$$KW = \frac{12}{32 \cdot 33} \left(\frac{157,5^2}{8} + \frac{213^2}{8} + \frac{116,5^2}{8} + \frac{41^2}{8} \right) - 3 \cdot 33 = 22,3473$$

Vzhledem k výskytu stejných údajů (bylo použito průměrné pořadí) je vhodné opravit testové kritérium korekčním faktorem:

$$K = 1 - \frac{(2^3 - 2) + (4^3 - 4) + (2^3 - 2) + (2^3 - 2) + (4^3 - 4) + (2^3 - 2)}{32^3 - 32} = 0,9956$$

$$KW_{opr.} = \frac{22,3473}{0,9956} = 22,446 \quad \chi^2_{0,05(3)} = 7,815 \quad \chi^2_{0,01(3)} = 11,34$$

2.3.2. Neparametrická analýza rozptylu

Příklad

Opravené testové kritérium $KW = 22,446 > \chi^2 \Rightarrow$ na hladině významnosti $\alpha = 0,05$ i $0,01$ přijímáme alternativní hypotézu, podle které se průkazně liší hodnoty výnosu nejméně ve 2 třídách.

Podrobnější vyhodnocení Neményiho metodou:

Spočteme tabulku diferencí mezi součty pořadí pro výnosy kukuřice při jednotlivých variantách hnojení $T_i - T_j$

2.3.2. Neparametrická analýza rozptylu

Příklad

Opravené testové kritérium $KW = 22,446 > \chi^2 \Rightarrow$ na hladině významnosti $\alpha = 0,05$ i $0,01$ přijímáme alternativní hypotézu, podle které se průkazně liší hodnoty výnosu nejméně ve 2 třídách.

Podrobnější vyhodnocení Neményiho metodou:

Spočteme tabulku diferencí mezi součty pořadí pro výnosy kukuřice při jednotlivých variantách hnojení $T_i - T_j$

třída $T_i \backslash$ třída T_j	diference $T_i - T_j$		
	V_2	V_3	V_4
V_1	55,5	41	116,5 ^x
V_2		96,5 ^x	172 ^{xx}
V_3			75,5

2.3.2. Neparametrická analýza rozptylu

Příklad

Opravené testové kritérium $KW = 22,446 > \chi^2 \Rightarrow$ na hladině významnosti $\alpha = 0,05$ i $0,01$ přijímáme alternativní hypotézu, podle které se průkazně liší hodnoty výnosu nejméně ve 2 třídách.

Podrobnější vyhodnocení Neményiho metodou:

Spočteme tabulku diferencí mezi součty pořadí pro výnosy kukuřice při jednotlivých variantách hnojení $T_i - T_j$

třída $T_i \backslash$ třída T_j	diference $T_i - T_j$		
	V_2	V_3	V_4
V_1	55,5	41	116,5 ^x
V_2		96,5 ^x	172 ^{xx}
V_3			75,5

Kritické hodnoty:

$$D_{0,05} = 96,4 \quad (N = 8, K = 4)$$

$$D_{0,01} = 116,8 \quad (N = 8, K = 4)$$

2.3.2. Neparametrická analýza rozptylu

Příklad

Opravené testové kritérium $KW = 22,446 > \chi^2 \Rightarrow$ na hladině významnosti $\alpha = 0,05$ i $0,01$ přijímáme alternativní hypotézu, podle které se průkazně liší hodnoty výnosu nejméně ve 2 třídách.

Podrobnější vyhodnocení Neményiho metodou:

Spočteme tabulku diferencí mezi součty pořadí pro výnosy kukuřice při jednotlivých variantách hnojení $T_i - T_j$

třída $T_i \backslash$ třída T_j	diference $T_i - T_j$		
	V_2	V_3	V_4
V_1	55,5	41	116,5 ^x
V_2		96,5 ^x	172 ^{xx}
V_3			75,5

Kritické hodnoty:

$$D_{0,05} = 96,4 \quad (N = 8, K = 4)$$

$$D_{0,01} = 116,8 \quad (N = 8, K = 4)$$

Statisticky významně se liší hodnoty výnosů kukuřice mezi variantami hnojení V_2, V_4 a to na hladině významnosti $\alpha = 0,01$ (diference $T_i - T_j$ jsou označeny ^{xx}). Na hladině významnosti $\alpha = 0,05$ se dále statisticky významně liší hodnoty výnosů kukuřice mezi variantami hnojení V_1, V_4 a V_2, V_3 .

2.3.2. Neparametrická analýza rozptylu

Test extrémních odchylek hodnot odezvy

V řadě pozorovaných hodnot se někdy objeví hodnota extrémně se lišící od ostatních, tzn. výrazně vybočuje z rozpětí ostatních naměřených hodnot.

Je třeba posoudit, zda je tato odchylka pouze náhodná nebo zda je uvedená hodnota zatížena „hrubou chybou“.

Pro objektivní posouzení této otázky existuje skupina testů, které se nazývají „*testy extrémních odchylek*“.

2.3.2. Neparametrická analýza rozptylu

Test extrémních odchylek hodnot odezvy

Dixonův test

Pozorovaná hodnota, která se extrémně liší od ostatních, je zřejmě buď nejmenší hodnotou ($x_{(1)}$) nebo největší hodnotou ($x_{(n)}$).

2.3.2. Neparametrická analýza rozptylu

Test extrémních odchylek hodnot odezvy

Dixonův test

Pozorovaná hodnota, která se extrémně liší od ostatních, je zřejmě buď nejmenší hodnotou ($x_{(1)}$) nebo největší hodnotou ($x_{(n)}$).

Nulová hypotéza H_0 tvrdí, že $(x_{(1)})$, resp. $(x_{(n)})$, je vybrána ze stejného normálně rozděleného základního souboru jako ostatní hodnoty.

Pro posouzení, zda hodnota $(x_{(1)})$ nebo hodnota $(x_{(n)})$ je zatížena hrubou chybou, užíváme testovacího kritéria

$$Q_1 = \frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}} \quad \text{nebo} \quad Q_n = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$$

2.3.2. Neparametrická analýza rozptylu

Test extrémních odchylek hodnot odezvy

Dixonův test

Pozorovaná hodnota, která se extrémně liší od ostatních, je zřejmě buď nejmenší hodnotou ($x_{(1)}$) nebo největší hodnotou ($x_{(n)}$).

Nulová hypotéza H_0 tvrdí, že ($x_{(1)}$), resp. ($x_{(n)}$), je vybrána ze stejného normálně rozděleného základního souboru jako ostatní hodnoty.

Pro posouzení, zda hodnota ($x_{(1)}$) nebo hodnota ($x_{(n)}$) je zatížena hrubou chybou, užíváme testovacího kritéria

$$Q_1 = \frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}} \quad \text{nebo} \quad Q_n = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$$

Jestliže vypočtená hodnota Q_1 , resp. Q_n , překročí kritickou hodnotu $Q_{1\alpha} = Q_{n\alpha}$ (nalezenou v tabulkách pro Dixonův test, hladinu významnosti α a rozsah souboru n), zamítáme nulovou hypotézu na hladině významnosti α a hodnotu ($x_{(1)}$), resp. ($x_{(n)}$), jako údaj zkreslený hrubou chybou, ze souboru vyloučíme.

Tzn. že nulová hypotéza se zamítá, pokud platí $Q_1 > Q_{\alpha(n)}$. nebo $Q_n > Q_{\alpha(n)}$.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Friedmanův test

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Friedmanův test

- Je neparametrickou obdobou analýzy rozptylu dvojného třídění (pro dva faktory - například pro jeden hlavní a druhý blokový).

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Friedmanův test

- Je neparametrickou obdobou analýzy rozptylu dvojného třídění (pro dva faktory - například pro jeden hlavní a druhý blokový).
- Předpokládejme, že máme měření y_{ij} , kde i je označení bloku, j je označení úrovně faktoru (ošetření). Pro každou dvojici (i,j) máme jedno měření.

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Friedmanův test

- Je neparametrickou obdobou analýzy rozptylu dvojného třídění (pro dva faktory - například pro jeden hlavní a druhý blokový).
- Předpokládejme, že máme měření y_{ij} , kde i je označení bloku, j je označení úrovně faktoru (ošetření). Pro každou dvojici (i,j) máme jedno měření.

- Označme

k - počet úrovní faktoru

n - počet bloků

r_{ij} - pořadí naměřené hodnoty odezvy y_{ij} v bloku i

$\bar{r}_j$ - průměrné pořadí odezvy při úrovni j přes všechny bloky

$\bar{r}$ - průměrné pořadí přes všechny úrovně a přes všechny bloky

$$\bar{r}_{\cdot j} = \frac{1}{n} \sum_{i=1}^n r_{ij} \quad \bar{r} = \frac{1}{nk} \sum_{i=1}^n \sum_{j=1}^k r_{ij}$$

2.3.2. Neparametrická analýza rozptylu

Odezva $X_1, X_2, \dots, X_n$ se neřídí normálním rozdělením => nelze použít ANOVA.

Friedmanův test

- Je neparametrickou obdobou analýzy rozptylu dvojného třídění (pro dva faktory - například pro jeden hlavní a druhý blokový).
- Předpokládejme, že máme měření y_{ij} , kde i je označení bloku, j je označení úrovně faktoru (ošetření). Pro každou dvojici (i,j) máme jedno měření.

- Označme

k - počet úrovní faktoru

n - počet bloků

r_{ij} - pořadí naměřené hodnoty odezvy y_{ij} v bloku i

$\bar{r}_{\cdot j}$ - průměrné pořadí odezvy při úrovni j přes všechny bloky

$\bar{r}$ - průměrné pořadí přes všechny úrovně a přes všechny bloky

$$\bar{r}_{\cdot j} = \frac{1}{n} \sum_{i=1}^n r_{ij} \quad \bar{r} = \frac{1}{nk} \sum_{i=1}^n \sum_{j=1}^k r_{ij}$$

Položme $SS_t = n \sum_{j=1}^k (\bar{r}_{\cdot j} - \bar{r})^2$ a $SS_e = \frac{1}{n(k-1)} \sum_{i=1}^n \sum_{j=1}^k (r_{ij} - \bar{r})^2$

Testová statistika má tvar $Q = \frac{SS_t}{SS_e}$

(Pro velká n a k ($n > 15$, $k > 4$) má přibližně chí-kvadrát rozdělení o $k-1$ stupních volnosti)

2.3.2. Neparametrická analýza rozptylu

The screenshot displays the Minitab software interface. The 'Stat' menu is open, showing the 'Nonparametrics' option highlighted. The 'Friedman' dialog box is open, showing the 'Response' field set to 'C1', 'Treatment' set to 'C3', and 'Blocks' set to 'C2'. The 'Store residuals' and 'Store fits' checkboxes are unchecked. The 'Select' button is visible at the bottom of the dialog box.

Session

Kruskal-Wallis Test

Kruskal-Wallis Test

C3	N	Median
1	2	16,7
2	2	16,7
3	2	17,7
4	2	16,7
5	2	16,7
Overall	10	5,5

H = 6,44 DF = 4 P = 0,169

* NOTE * One or more small samples

Friedman Test: C1 versus C3 blocked b

S = 1,50 DF = 2 P = 0,472

C3	N	Est Median	Sum of Ranks
1	4	16,747	9,0
2	4	16,460	6,0
3	4	16,778	9,0

Grand median = 16,662

Friedman

Response: C1

Treatment: C3

Blocks: C2

☐ Store residuals

☐ Store fits

Select

Help OK Cancel

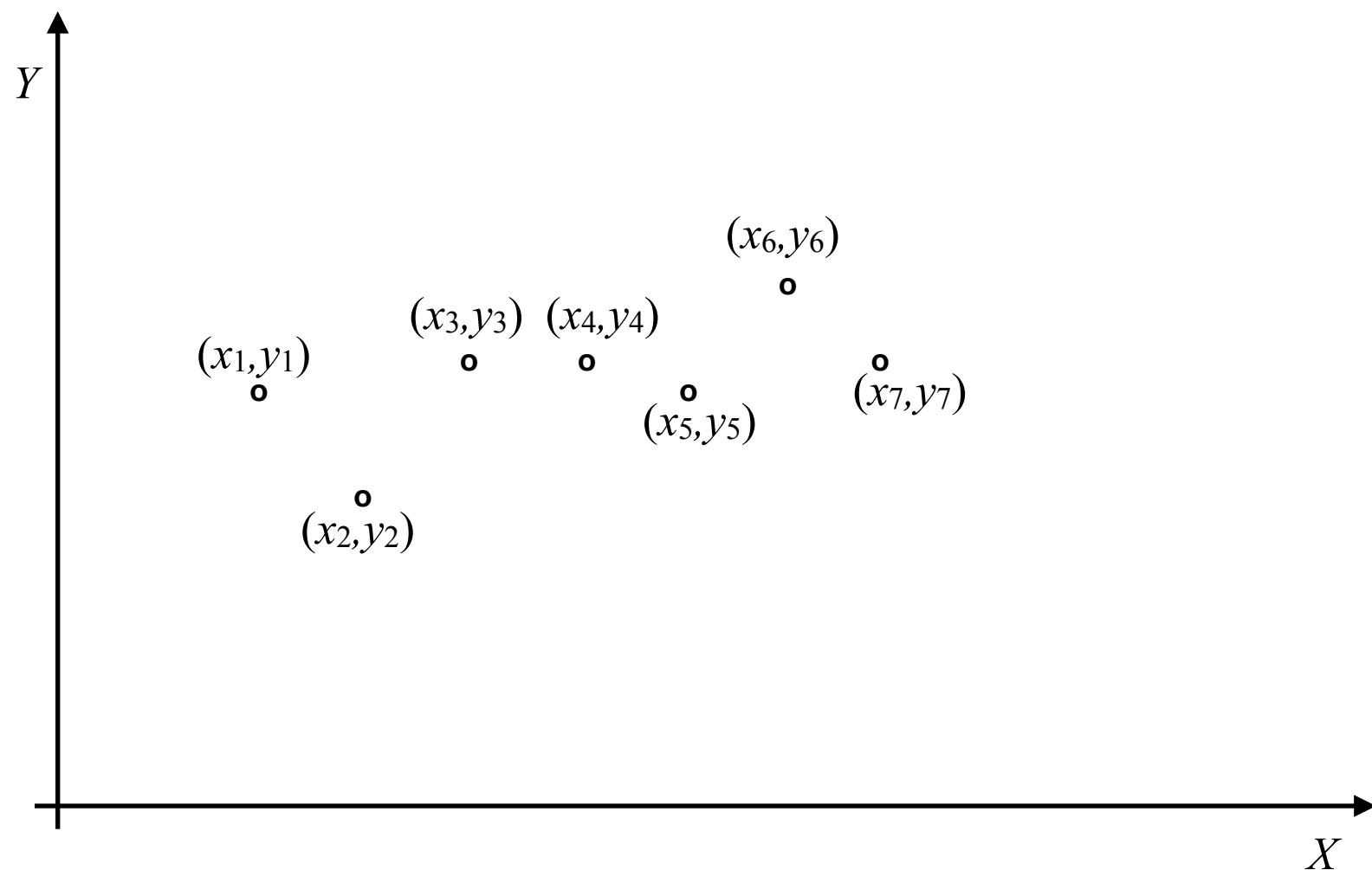
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .



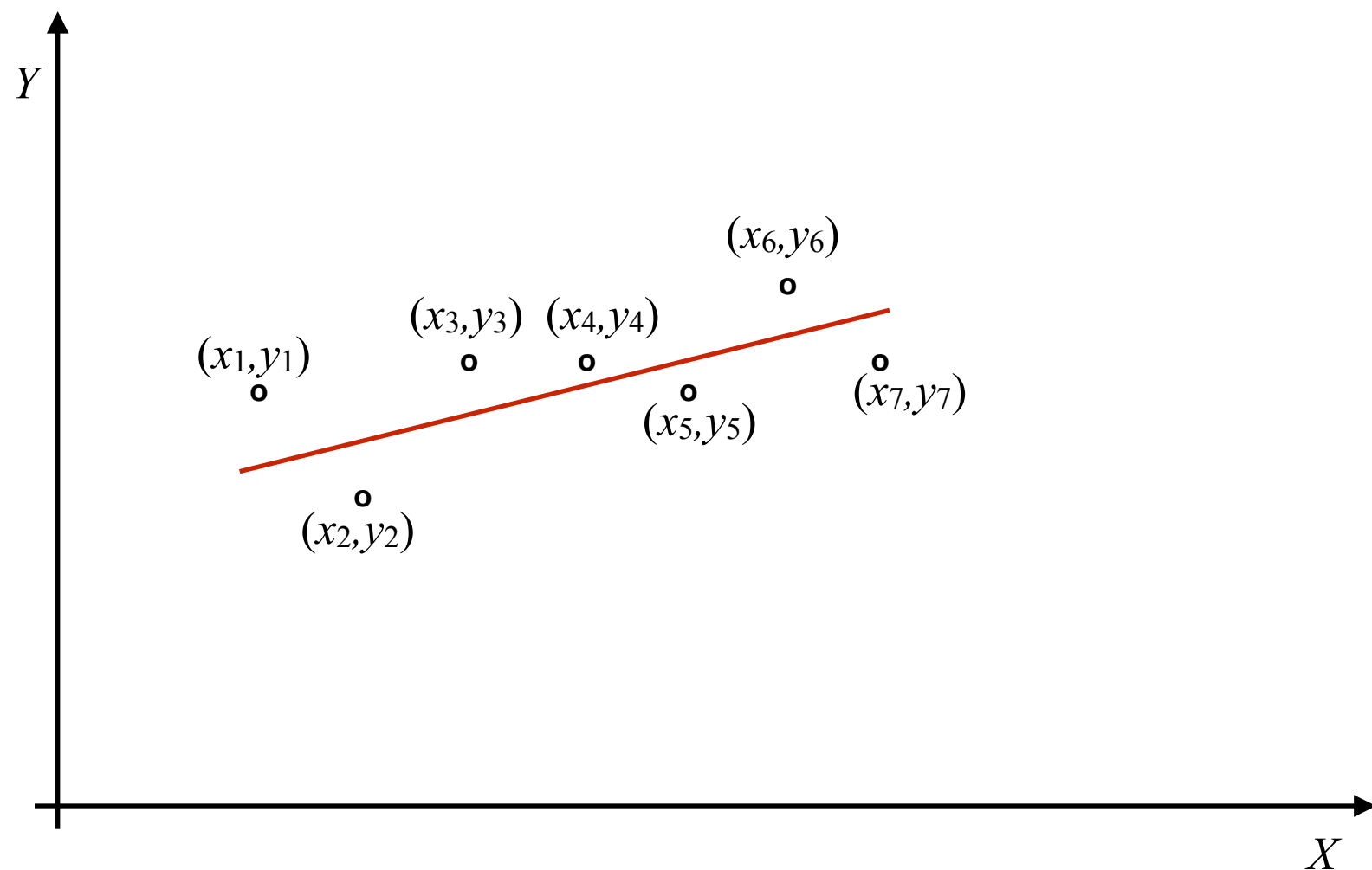
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .



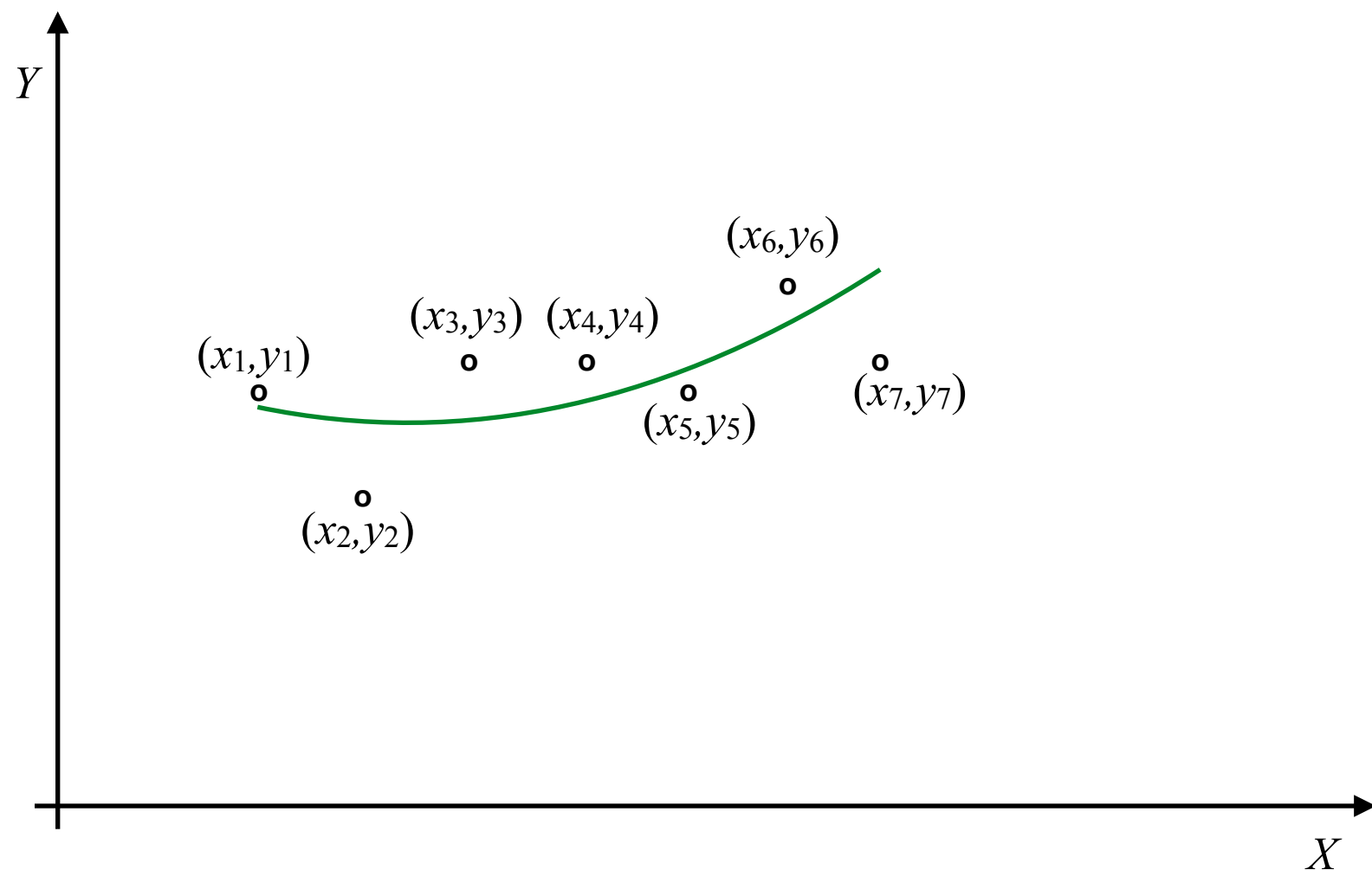
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + cX^2 + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .



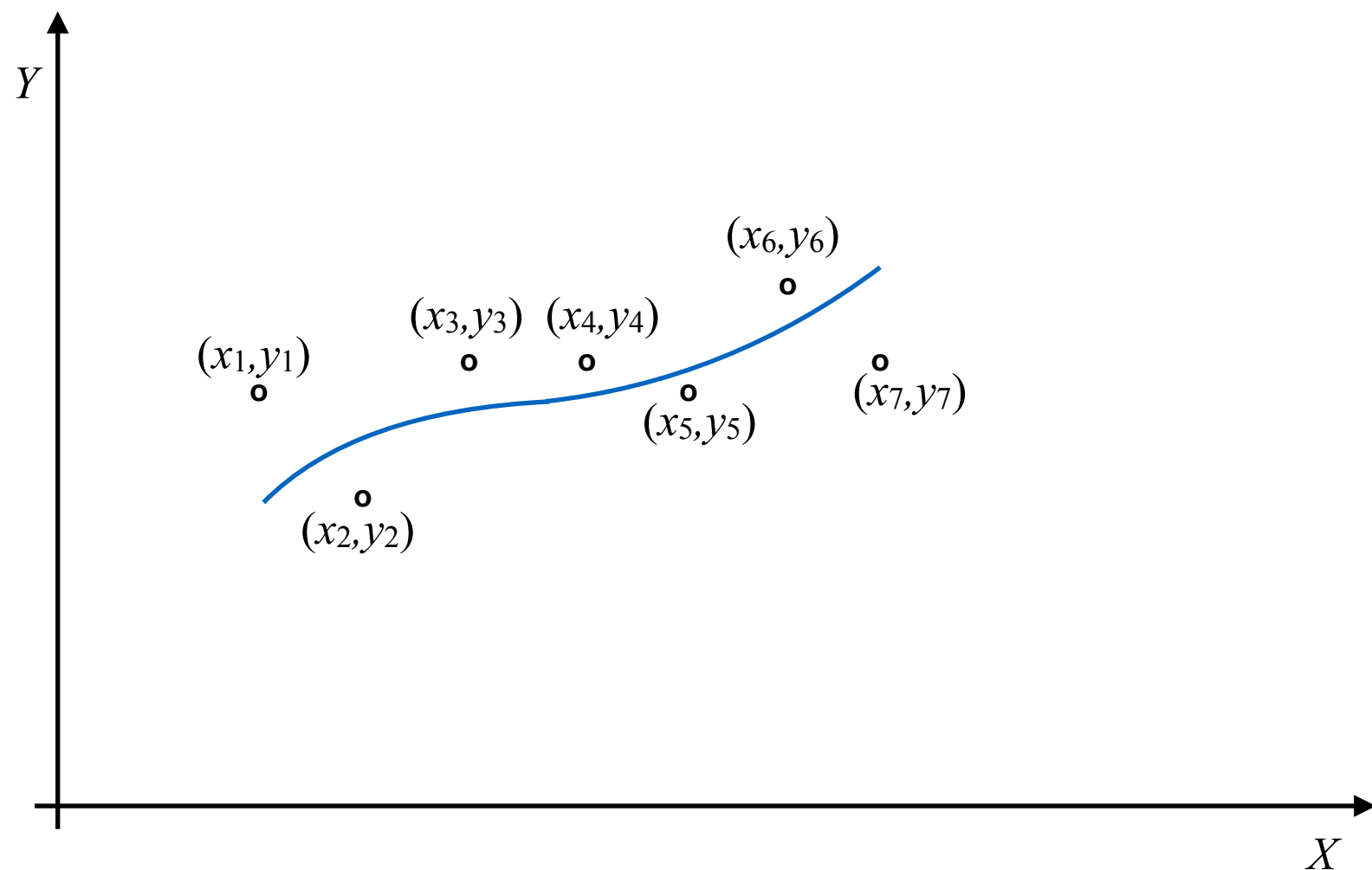
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + cX^2 + dX^3 + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .



2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

1) Odhadneme parametry regresní funkce (metoda nejmenších čtverců):

Označíme: $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix}$ $\mathbf{X} = \begin{pmatrix} 1 & X_1 \\ \dots & \dots \\ 1 & X_n \end{pmatrix}$ $\boldsymbol{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}$ $\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$ $\mathbf{b} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$

Odhad parametrů metodou nejmenších čtverců: $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{Y} \Rightarrow \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

1) Odhadneme parametry regresní funkce (metoda nejmenších čtverců):

Označíme: $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix}$ $\mathbf{X} = \begin{pmatrix} 1 & X_1 \\ \dots & \dots \\ 1 & X_n \end{pmatrix}$ $\boldsymbol{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}$ $\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$ $\mathbf{b} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$

Odhad parametrů metodou nejmenších čtverců: $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{Y} \Rightarrow \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

$$\hat{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}, \quad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

1) Odhadneme parametry regresní funkce (metoda nejmenších čtverců):

Označíme: $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix}$ $\mathbf{X} = \begin{pmatrix} 1 & X_1 \\ \dots & \dots \\ 1 & X_n \end{pmatrix}$ $\boldsymbol{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}$ $\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$ $\mathbf{b} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$

Odhad parametrů metodou nejmenších čtverců: $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{Y} \Rightarrow \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

$$\hat{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}, \quad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

Platí: $E(\mathbf{b}) = \boldsymbol{\beta}$, $D(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

1) Odhadneme parametry regresní funkce (metoda nejmenších čtverců):

Označíme: $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix}$ $\mathbf{X} = \begin{pmatrix} 1 & X_1 \\ \dots & \dots \\ 1 & X_n \end{pmatrix}$ $\boldsymbol{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}$ $\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$ $\mathbf{b} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$

Odhad parametrů metodou nejmenších čtverců: $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{Y} \Rightarrow \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

$$\hat{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}, \quad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

Platí: $E(\mathbf{b}) = \boldsymbol{\beta}$, $D(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

2) Spočteme koeficient determinace - míru těsnosti regrese: $R^2 = \frac{\sum (\hat{a} + \hat{b}x_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$

Koeficient determinace by měl být blízko 1 (alespoň $> 0,6$)

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

3) Provedeme analýzu reziduí: označme $\hat{y}_i = \hat{a} - \hat{b}x_i$, potom $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$

Měli bychom otestovat, zda rezidua e_i odpovídají předpokladu normálního rozdělení $N(0, \sigma^2)$ s nulovou střední hodnotou a rozptylem nezávislým na indexu i (test normality)

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

3) Provedeme analýzu reziduí: označme $\hat{y}_i = \hat{a} - \hat{b}x_i$, potom $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$

Měli bychom otestovat, zda rezidua e_i odpovídají předpokladu normálního rozdělení $N(0, \sigma^2)$ s nulovou střední hodnotou a rozptylem nezávislým na indexu i (test normality)

4) Provedeme test významnosti regresních koeficientů:

Spočteme odhad reziduálního rozptylu $s^2 = \frac{\sum y_i^2 - \hat{a} \sum y_i - \hat{b} \sum x_i y_i}{n - 2}$

Pro test hypotézy $H_0: b = 0$ použijeme testovou statistiku $T_b = \frac{\hat{b}}{s} \sqrt{\sum x_i^2 - n\bar{x}^2}$, která má t-rozdělení o $(n-2)$ stupních volnosti.

2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

Experiment: Provedeme n měření odezvy Y při n různých hodnotách sledovaného faktoru X .

3) Provedeme analýzu reziduí: označme $\hat{y}_i = \hat{a} - \hat{b}x_i$, potom $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$

Měli bychom otestovat, zda rezidua e_i odpovídají předpokladu normálního rozdělení $N(0, \sigma^2)$ s nulovou střední hodnotou a rozptylem nezávislým na indexu i (test normality)

4) Provedeme test významnosti regresních koeficientů:

Spočteme odhad reziduálního rozptylu
$$s^2 = \frac{\sum y_i^2 - \hat{a} \sum y_i - \hat{b} \sum x_i y_i}{n - 2}$$

Pro test hypotézy $H_0: b = 0$ použijeme testovou statistiku
$$T_b = \frac{\hat{b}}{s} \sqrt{\sum x_i^2 - n\bar{x}^2},$$
 která má t-rozdělení o $(n-2)$ stupních volnosti.

5) Spočteme interval spolehlivosti pro predikci:

Pro dané x je interval spolehlivosti predikce Y :
$$\hat{a} + \hat{b}x \pm t_{n-2}(\alpha) s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum x_i^2 - n\bar{x}^2}}$$

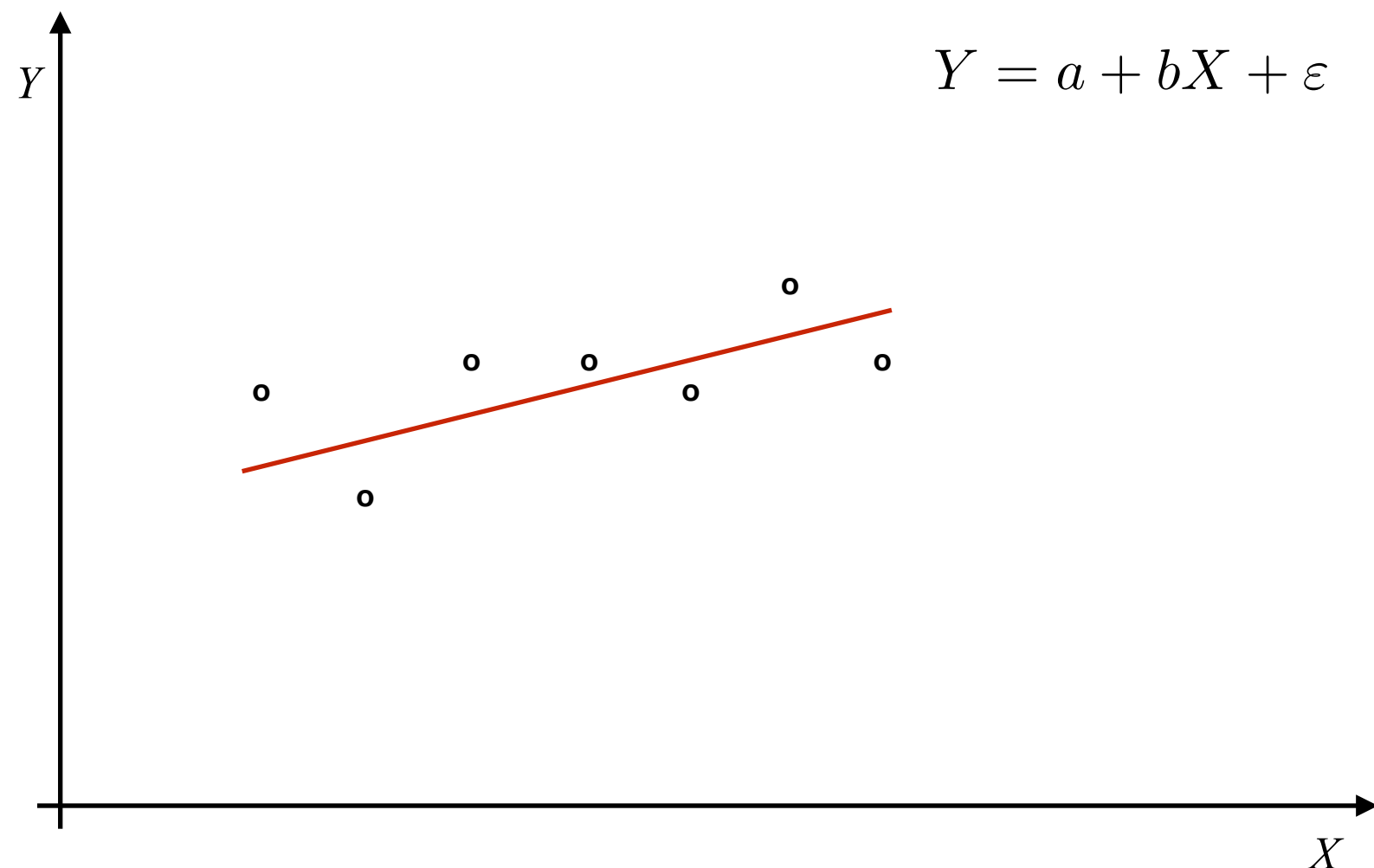
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

5) Spočteme pásy spolehlivosti:



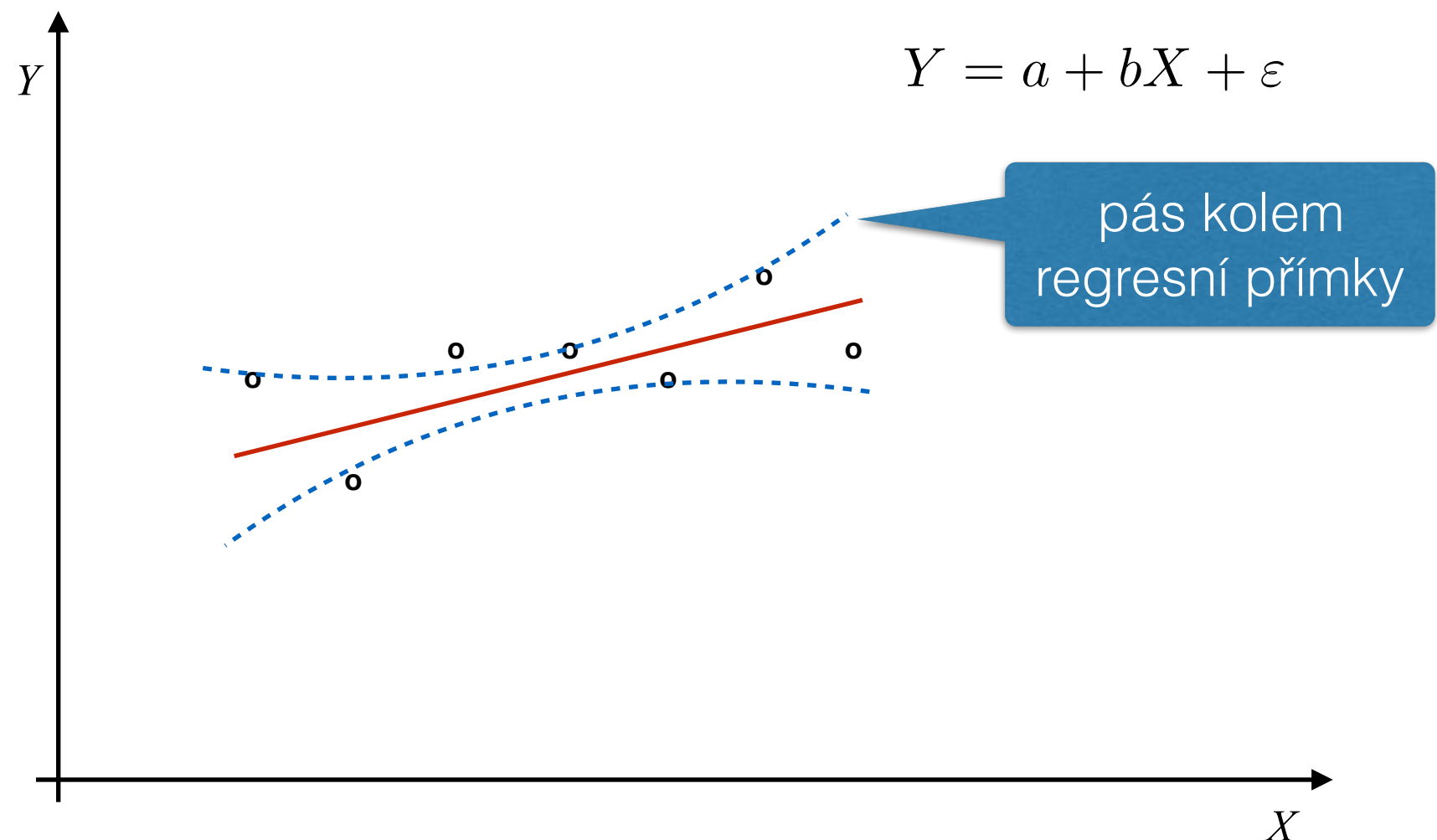
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

5) Spočteme pásy spolehlivosti:



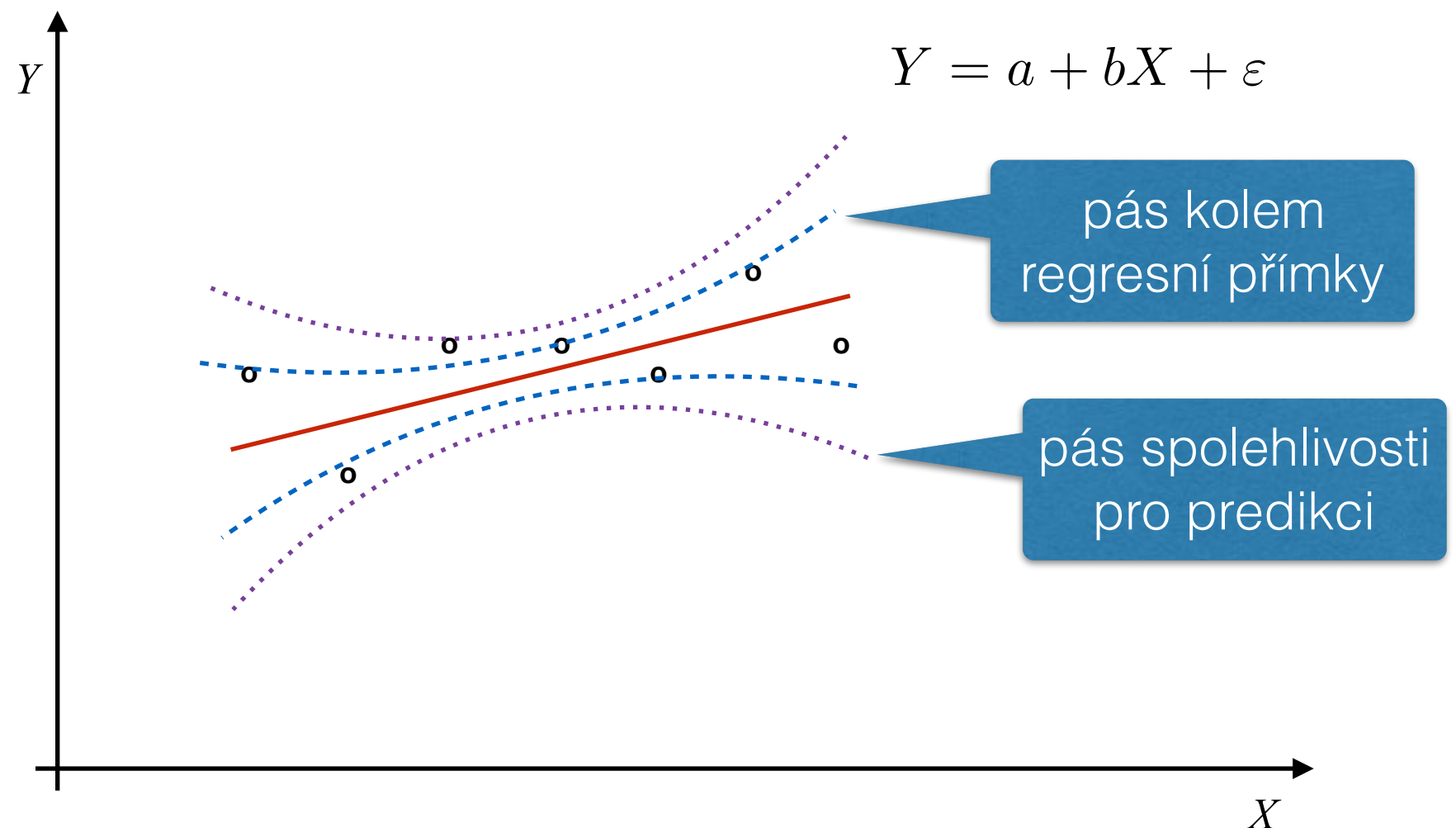
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

5) Spočteme pásy spolehlivosti:



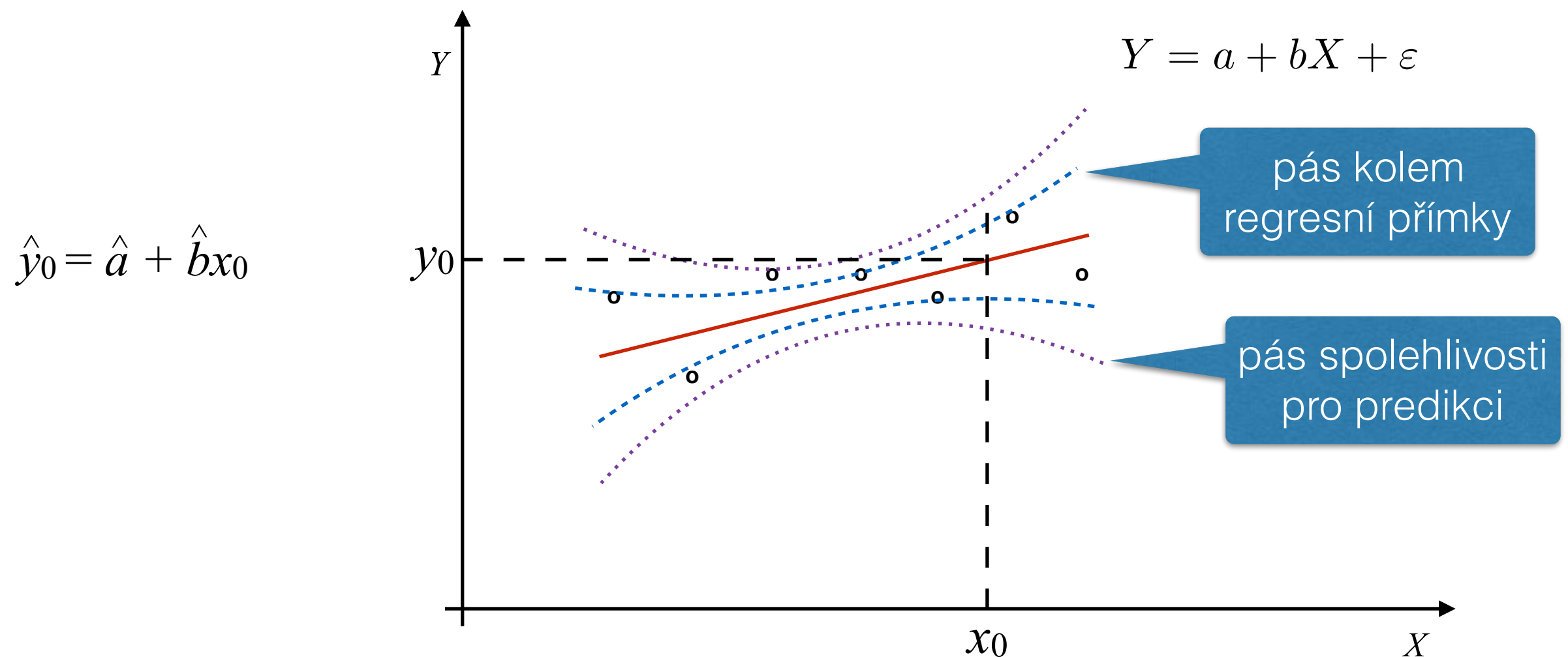
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

5) Spočteme pásy spolehlivosti:



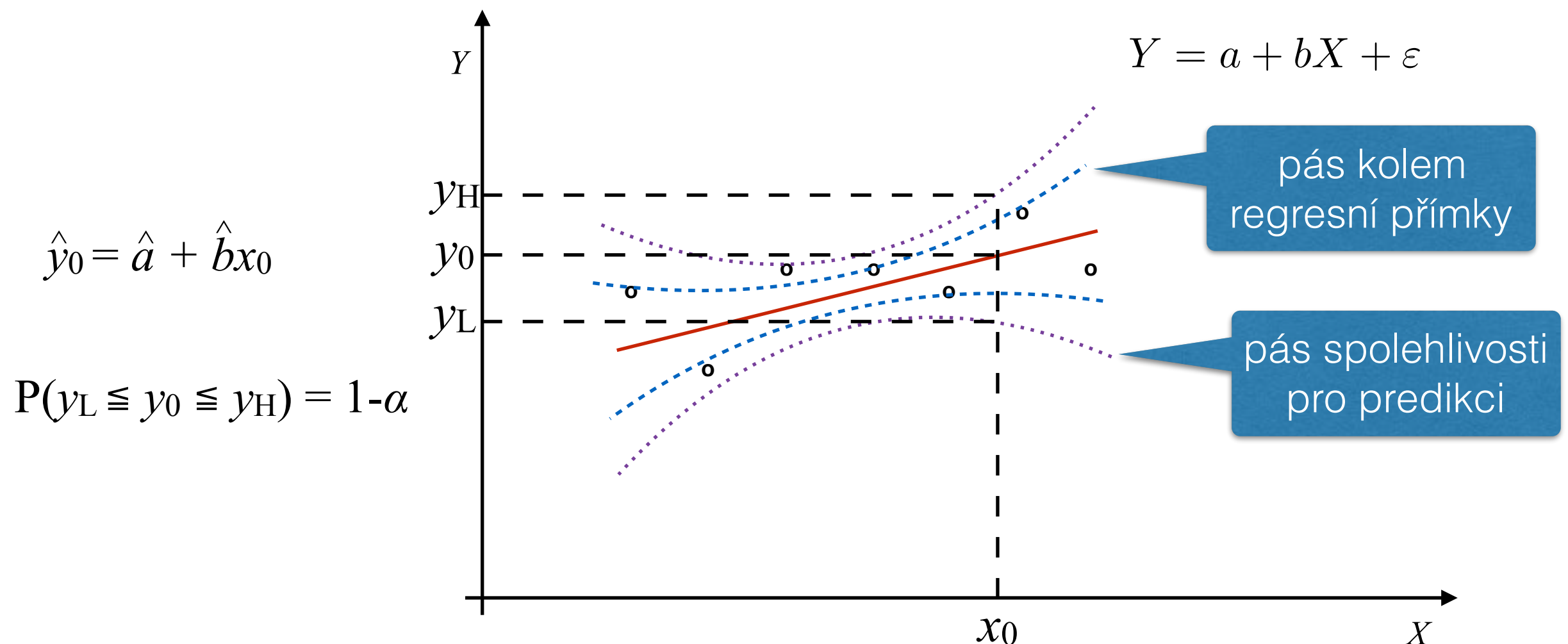
2.4. Regresní experiment pro jeden faktor o n úrovních

Úloha: Máme najít závislost mezi spojitou odezvou Y a hodnotami vybraného faktoru X .

Lineární model pro jeden faktor:

$$Y = a + bX + \varepsilon$$

5) Spočteme pásy spolehlivosti:



2.4. Regresní analýza

Regresní model experimentu

lineární model pro dva faktory:

$$Y = a + bX + cZ + dXZ + \varepsilon$$

2.4. Regresní analýza

Regresní model experimentu

lineární model pro dva faktory:

$$Y = a + bX + cZ + dXZ + \varepsilon$$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & X_1 & Z_1 & X_1 Z_1 \\ \dots & \dots & \dots & \dots \\ 1 & X_n & Z_n & X_n Z_n \end{pmatrix} \quad \beta = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$$

2.4. Regresní analýza

Regresní model experimentu

lineární model pro dva faktory:

$$Y = a + bX + cZ + dXZ + \varepsilon$$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & X_1 & Z_1 & X_1 Z_1 \\ \dots & \dots & \dots & \dots \\ 1 & X_n & Z_n & X_n Z_n \end{pmatrix} \quad \beta = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$$
$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \sum x_i & \sum z_i & \sum x_i z_i \\ \sum x_i & \sum x_i^2 & \sum x_i z_i & \sum x_i^2 z_i \\ \sum z_i & \sum x_i z_i & \sum z_i^2 & \sum x_i z_i^2 \\ \sum x_i z_i & \sum x_i^2 z_i & \sum x_i z_i^2 & \sum x_i^2 z_i^2 \end{pmatrix} \quad \mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum Y_i \\ \sum x_i Y_i \\ \sum z_i Y_i \\ \sum x_i z_i Y_i \end{pmatrix}$$

2.4. Regresní analýza

Regresní model experimentu

lineární model pro dva faktory:

$$Y = a + bX + cZ + dXZ + \varepsilon$$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & X_1 & Z_1 & X_1 Z_1 \\ \dots & \dots & \dots & \dots \\ 1 & X_n & Z_n & X_n Z_n \end{pmatrix} \quad \beta = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$$
$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \sum x_i & \sum z_i & \sum x_i z_i \\ \sum x_i & \sum x_i^2 & \sum x_i z_i & \sum x_i^2 z_i \\ \sum z_i & \sum x_i z_i & \sum z_i^2 & \sum x_i z_i^2 \\ \sum x_i z_i & \sum x_i^2 z_i & \sum x_i z_i^2 & \sum x_i^2 z_i^2 \end{pmatrix} \quad \mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum Y_i \\ \sum x_i Y_i \\ \sum z_i Y_i \\ \sum x_i z_i Y_i \end{pmatrix}$$

