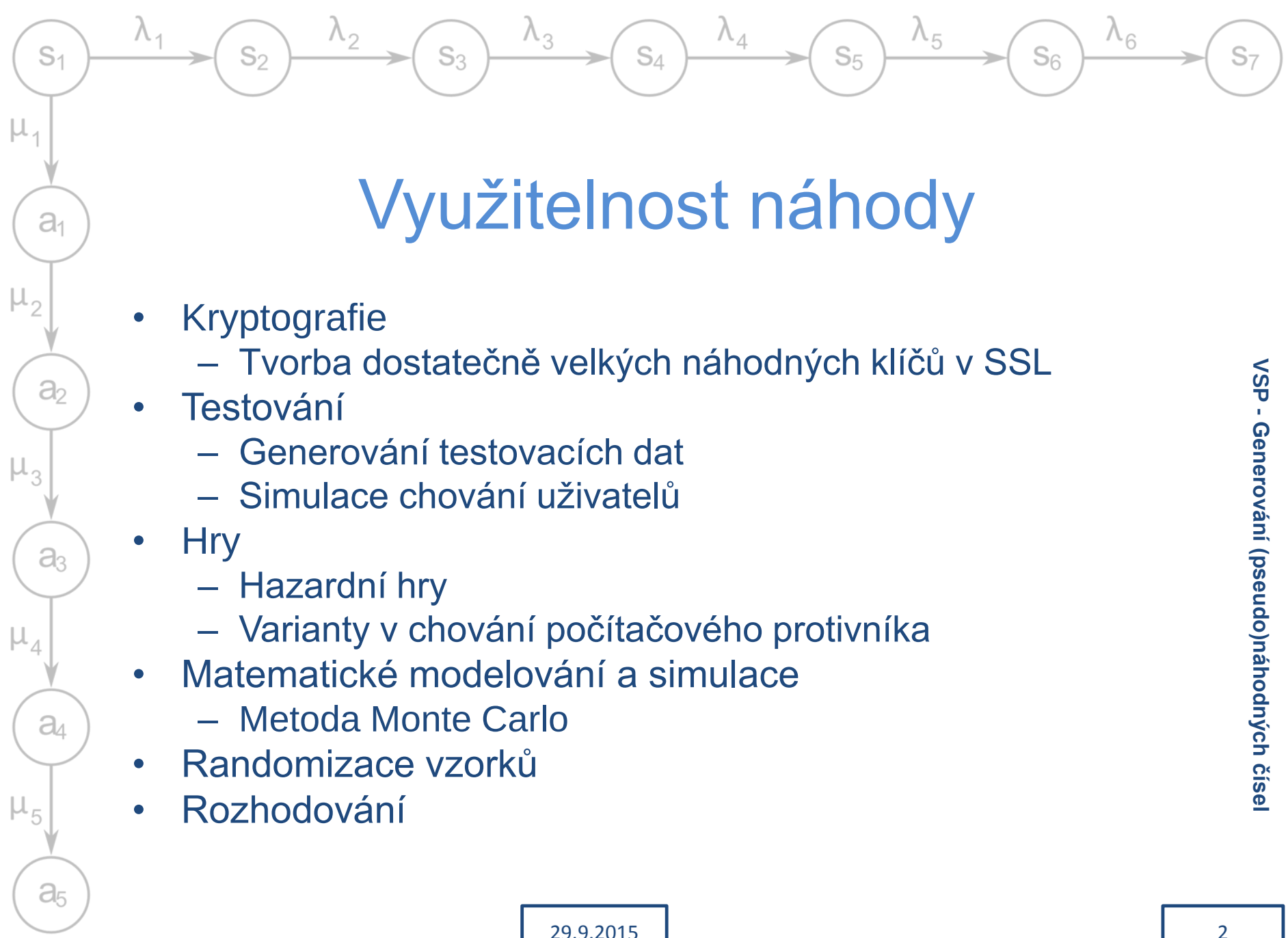
Generování (pseudo)náhodných čísel

Výkonnost a spolehlivost – KIV/VSP

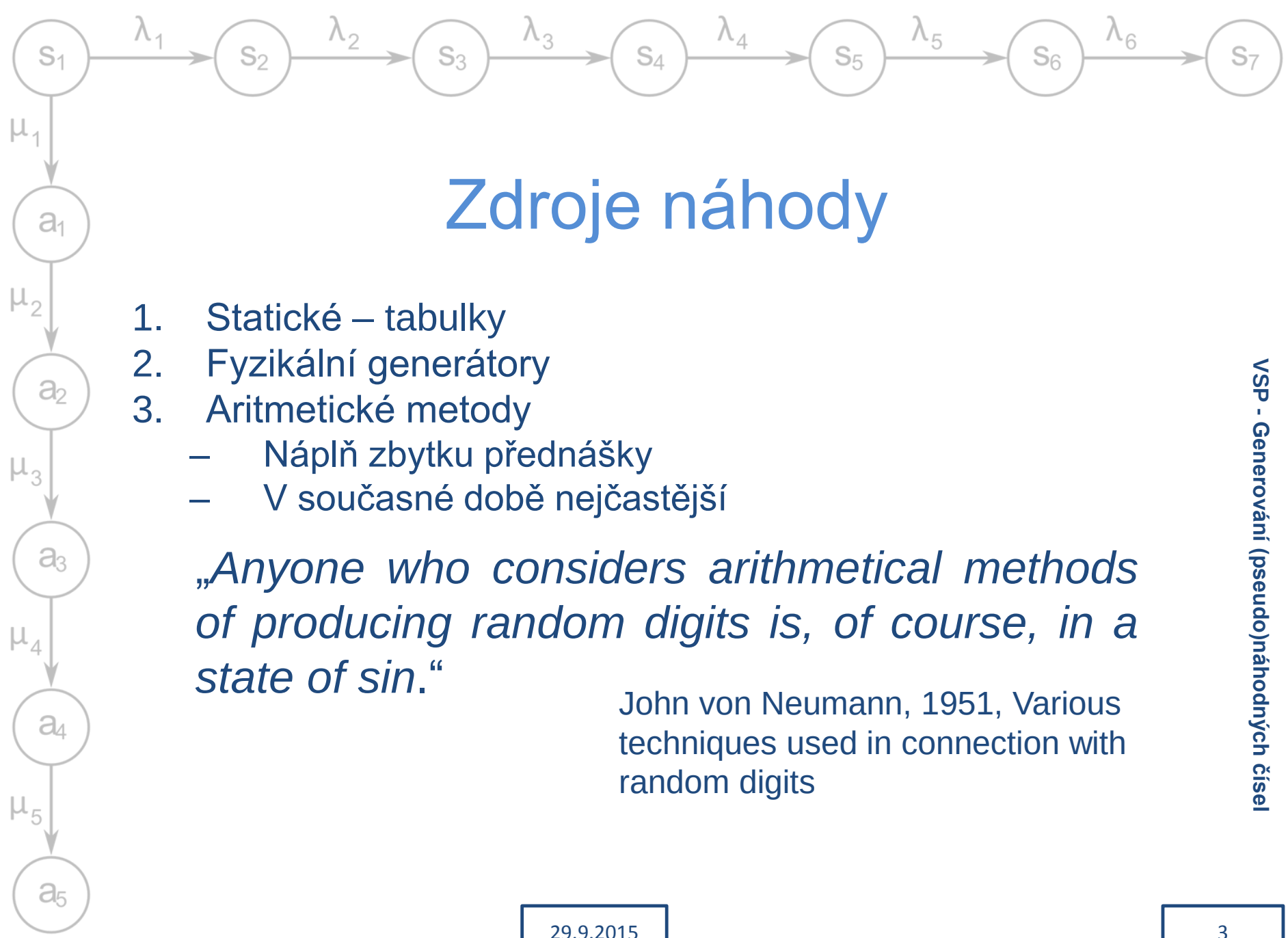
Richard Lipka

29.9.2015



Využitelnost náhody

- Kryptografie
  - Tvorba dostatečně velkých náhodných klíčů v SSL
- Testování
  - Generování testovacích dat
  - Simulace chování uživatelů
- Hry
  - Hazardní hry
  - Varianty v chování počítačového protivníka
- Matematické modelování a simulace
  - Metoda Monte Carlo
- Randomizace vzorků
- Rozhodování

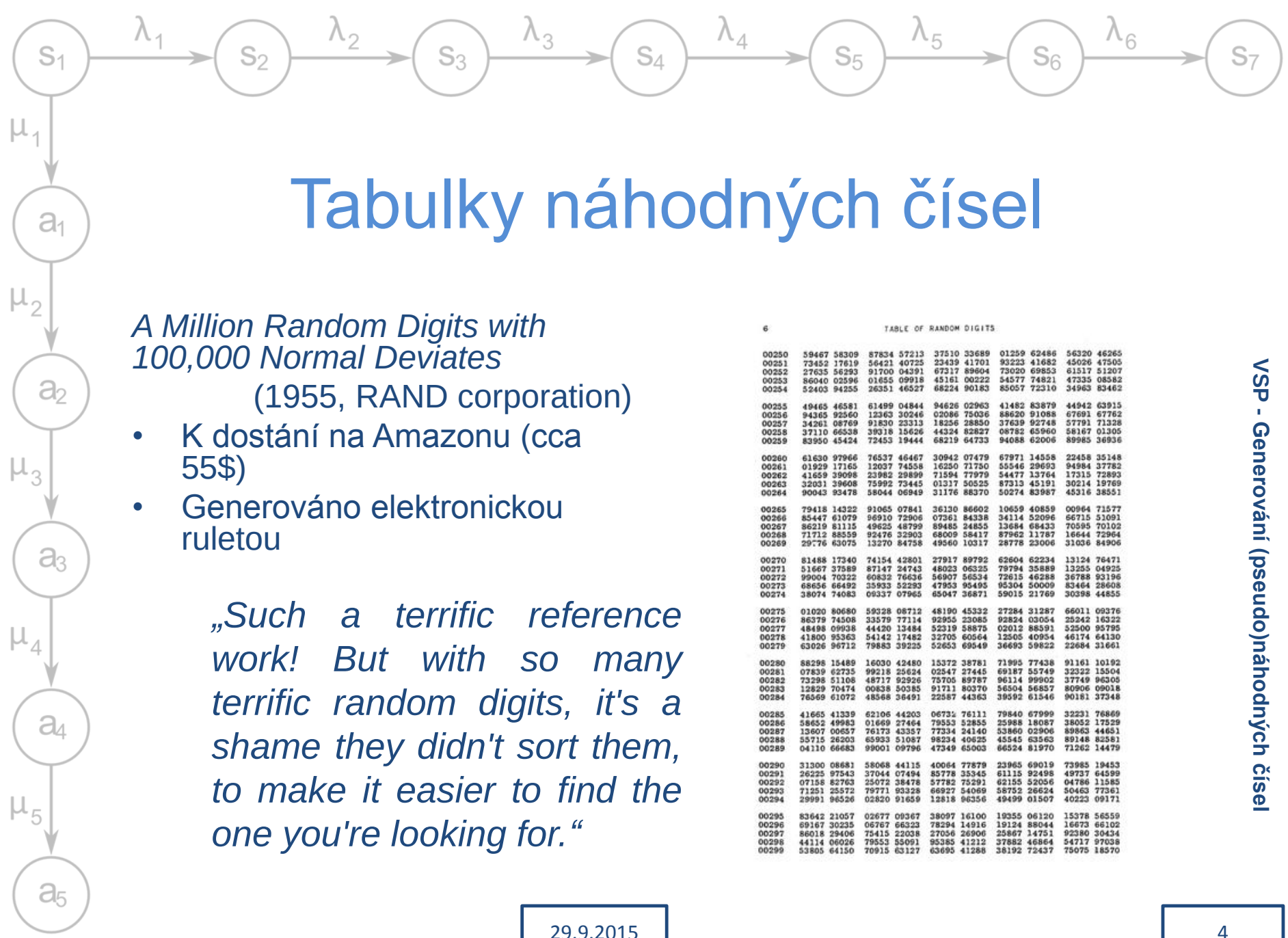


Zdroje náhody

1. Statické – tabulky
2. Fyzikální generátory
3. Aritmetické metody
   - Náplň zbytku přednášky
   - V současné době nejčastější

„*Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin*.“

John von Neumann, 1951, Various techniques used in connection with random digits



Tabulky náhodných čísel

A Million Random Digits with 100,000 Normal Deviates

(1955, RAND corporation)

- K dostání na Amazonu (cca 55$)
- Generováno elektronickou ruletou

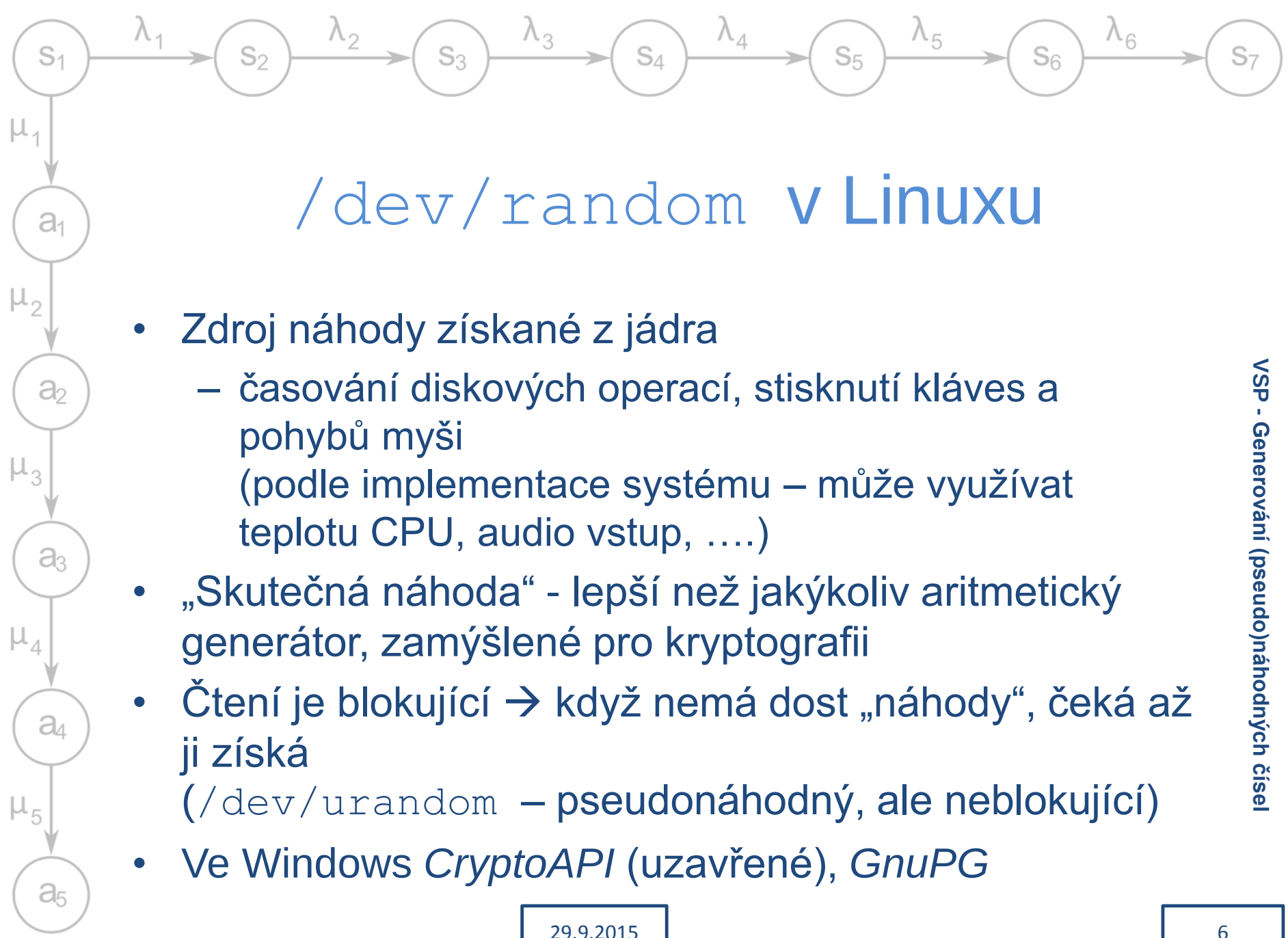
„Such a terrific reference work! But with so many terrific random digits, it's a shame they didn't sort them, to make it easier to find the one you're looking for.“

6 TABLE OF RANDOM DIGITS

00250	59467	58309	87834	57213	37510	33689	01259	62486	56320	46265
00251	73452	17619	56421	40725	23439	41701	93223	41682	45026	47505
00252	27635	56293	91700	04391	67317	89604	73020	69853	61517	51207
00253	86040	02596	01655	09918	45161	00222	54577	74821	47335	08582
00254	52403	94255	26351	46527	68224	90183	85057	72310	34963	83462
00255	49465	46581	61499	04844	94626	02963	41482	83879	44942	63915
00256	94365	92560	12363	30246	02086	75036	88620	91088	67691	67762
00257	34261	08769	91830	23313	18256	28850	37639	92748	57791	71328
00258	37110	66538	39318	15626	44324	82827	08782	65960	58167	01305
00259	83950	45424	72453	19444	68219	64733	94088	62006	89985	36936
00260	61630	97966	76537	46467	30942	07479	67971	14558	22458	35148
00261	01929	17165	12037	74558	16250	71750	55546	29693	94984	37782
00262	41659	39098	23982	29899	71594	77979	54477	13764	17315	72893
00263	32031	39608	75992	73445	01317	50525	87313	45191	30214	19769
00264	90043	93478	58044	06949	31176	88370	50274	83987	45316	38551
00265	79418	14322	91065	07841	36130	86602	10659	40859	00964	71577
00266	85447	61079	96910	72906	07361	84338	34114	52096	66715	51091
00267	86219	81115	49625	48799	89485	24855	13684	68433	70595	70102
00268	71712	88559	92476	32903	68009	58417	87962	11787	16644	72964
00269	29776	63075	13270	84758	49560	10317	28778	23006	31036	84906
00270	81488	17340	74154	42801	27917	89792	62604	62234	13124	76471
00271	51667	37589	87147	24743	48023	06325	79794	35889	13255	04925
00272	99004	70322	60832	76636	56907	56534	72615	46288	36788	93196
00273	68656	66492	35933	52293	47953	95495	95304	50009	83464	28608
00274	38074	74083	09337	07965	65047	36871	59015	21769	30398	44855
00275	01020	80680	59328	08712	48190	45332	27284	31287	66011	09376
00276	86379	74508	33579	77114	92955	23085	92824	03054	25242	16322
00277	48498	09938	44420	13484	52319	58875	02012	88591	52500	95795
00278	41800	95363	54142	17482	32705	60564	12505	40954	46174	64130
00279	63026	96712	79883	39225	52653	69549	36693	59822	22684	31661
00280	88298	15489	16030	42480	15372	38781	71995	77438	91161	10192
00281	07839	62735	99218	25624	02547	27445	69187	55749	32322	15504
00282	73298	51108	48717	92926	75705	89787	96114	99902	37749	96305
00283	12829	70474	00838	50385	91711	80370	56504	56857	80906	09018
00284	76569	61072	48568	36491	22587	44363	39592	61546	90181	37348
00285	41665	41339	62106	44203	06734	76111	79840	67999	32231	76869
00286	58652	49983	01669	27464	79553	52855	25988	18087	38052	17529
00287	13607	00657	76173	43357	77334	24140	53860	02906	89863	44651
00288	55715	26203	65933	51087	98234	40625	45545	63563	89148	82581
00289	04110	66683	99001	09796	47349	65003	66524	81970	71262	14479
00290	31300	08681	58068	44115	40064	77879	23965	69019	73985	19453
00291	26225	97543	37044	07494	85778	35345	61115	92498	49737	64599
00292	07158	82763	25072	38478	57782	75291	62155	52056	04786	11585
00293	71251	25572	79771	93328	66927	54069	58752	26624	50463	77361
00294	29991	96526	02820	91659	12818	96356	49499	01507	40223	09171
00295	83642	21057	02677	09367	38097	16100	19355	06120	15378	56559
00296	69167	30235	06767	66323	78294	14916	19124	88044	16673	66102
00297	86018	29406	75415	22038	27056	26906	25867	14751	92380	30434
00298	44114	06026	79553	55091	95385	41212	37882	46864	54717	97038
00299	53805	64150	70915	63127	63695	41288	38192	72437	75075	18570

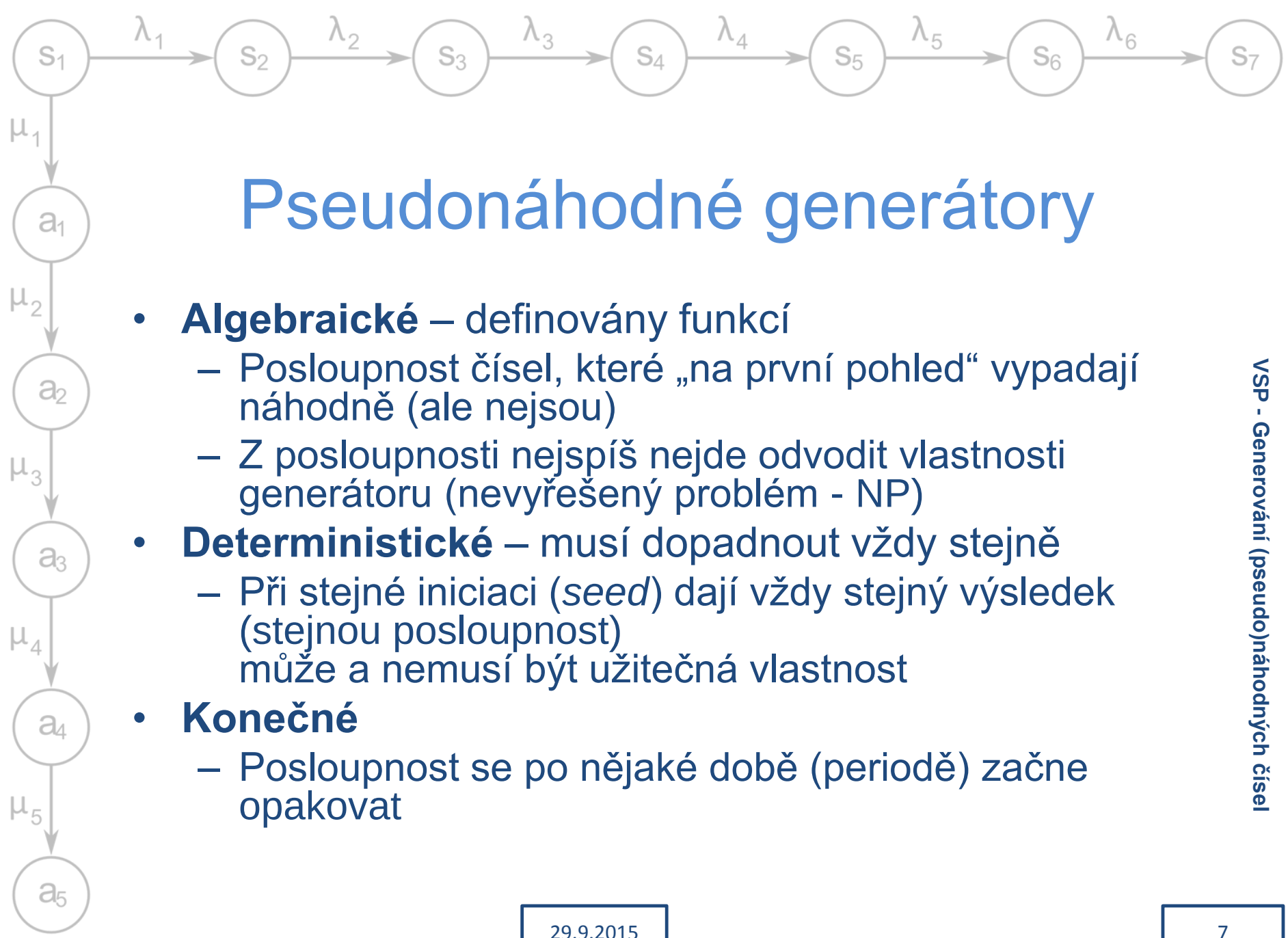
Fyzikální generátory

- Náhoda získána měřením nějakého reálného procesu
  - Čítače částic
  - Tepelný šum, elektromagnetický šum
  - Počasí
  - Házení kostek, míchání karet
  - …
- Jevy které je velmi obtížné předvídat
  → využitelné v kryptografii



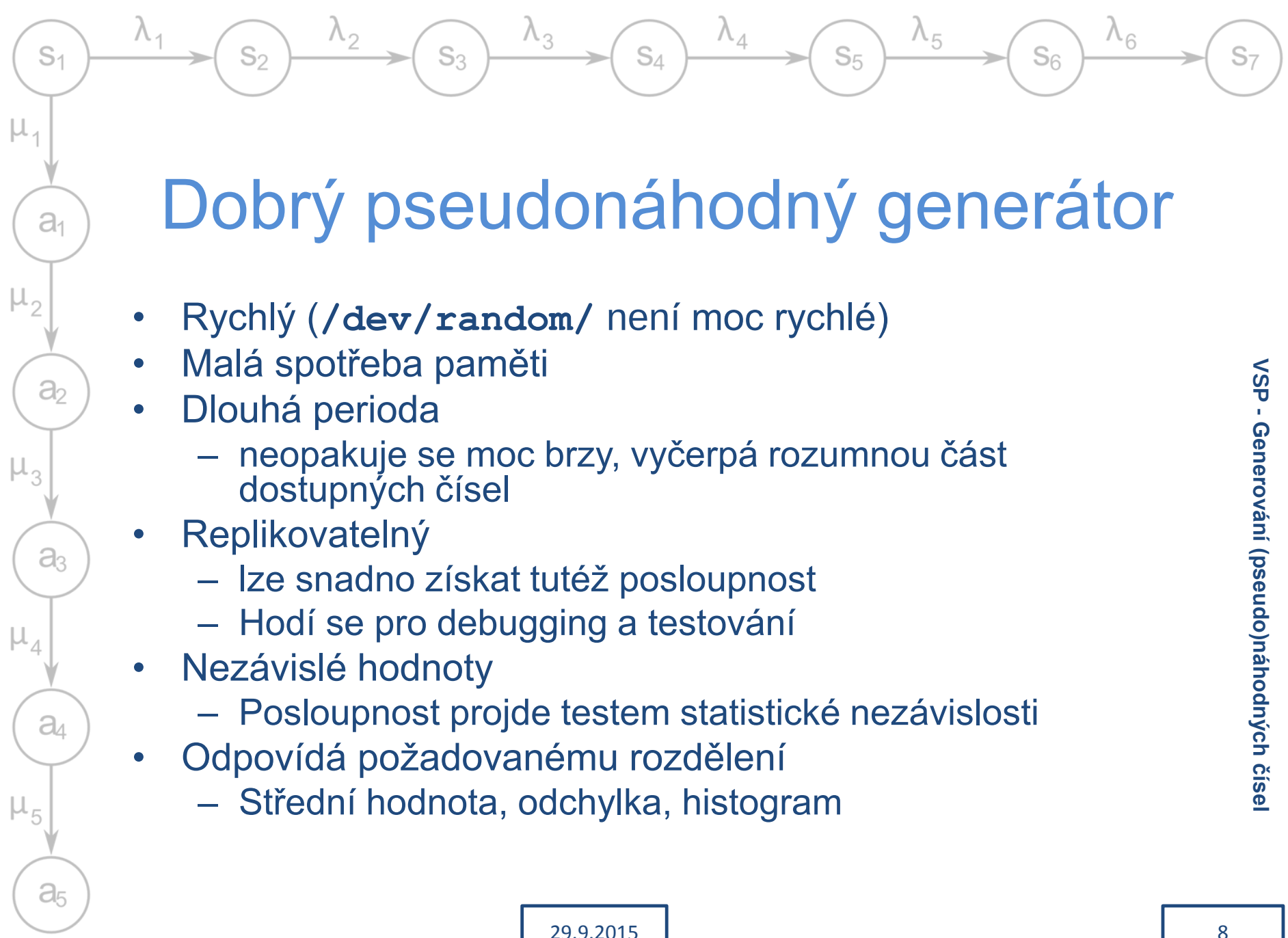
/dev/random v Linuxu

- Zdroj náhody získané z jádra
  - časování diskových operací, stisknutí kláves a pohybů myši
    (podle implementace systému – může využívat teplotu CPU, audio vstup, ….)
- „Skutečná náhoda“ - lepší než jakýkoliv aritmetický generátor, zamýšlené pro kryptografii
- Čtení je blokující → když nemá dost „náhody“, čeká až ji získá
  (`/dev/urandom` – pseudonáhodný, ale neblokující)
- Ve Windows *CryptoAPI* (uzavřené), *GnuPG*



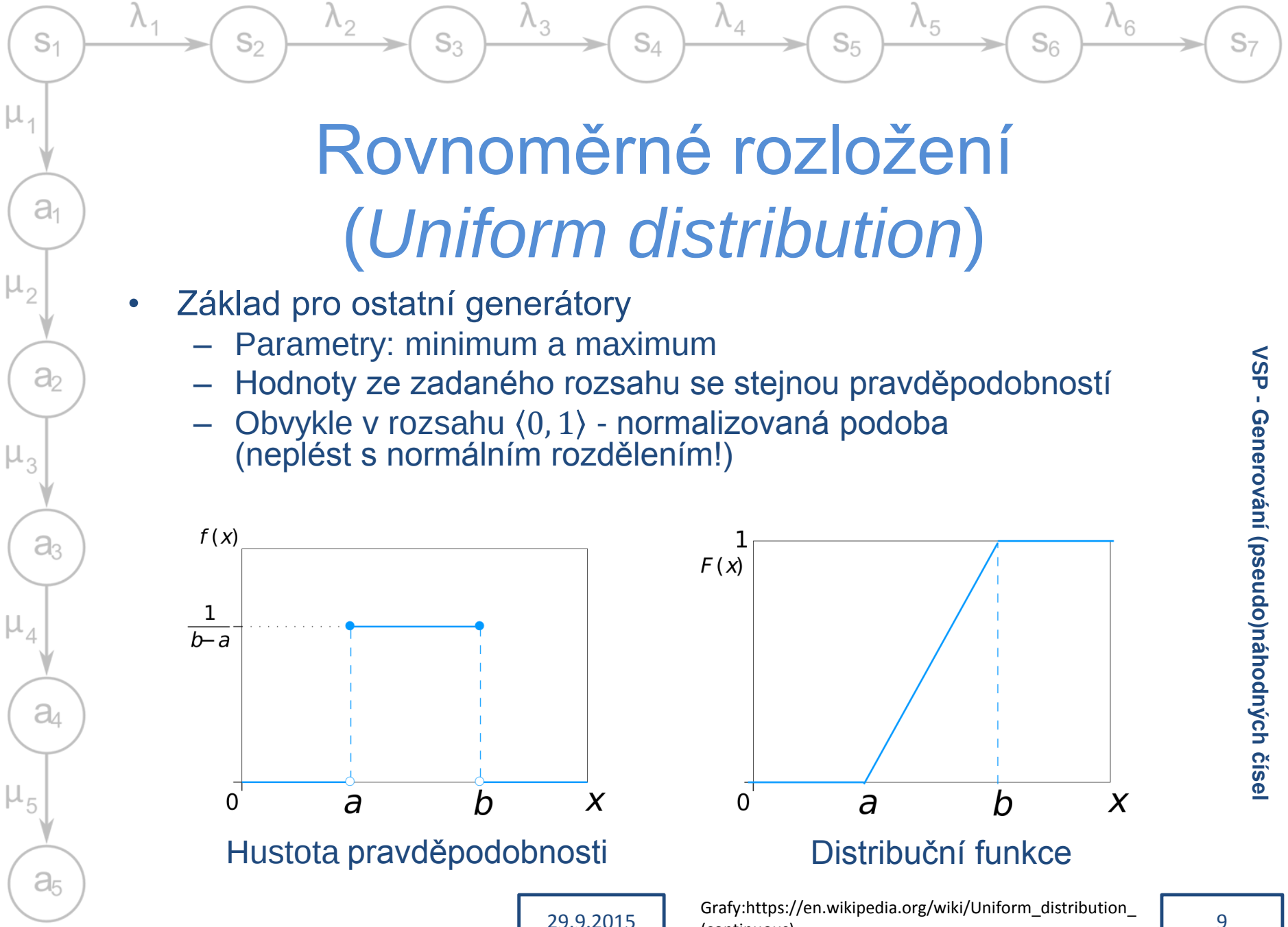
Pseudonáhodné generátory

- **Algebraické** – definovány funkcí
  - Posloupnost čísel, které „na první pohled“ vypadají náhodně (ale nejsou)
  - Z posloupnosti nejspíš nejde odvodit vlastnosti generátoru (nevyřešený problém - NP)
- **Deterministické** – musí dopadnout vždy stejně
  - Při stejné iniciaci (*seed*) dají vždy stejný výsledek (stejnou posloupnost)
    může a nemusí být užitečná vlastnost
- **Konečné**
  - Posloupnost se po nějaké době (periodě) začne opakovat



Dobrý pseudonáhodný generátor

- Rychlý (**`/dev/random/`** není moc rychlé)
- Malá spotřeba paměti
- Dlouhá perioda
  - neopakuje se moc brzy, vyčerpá rozumnou část dostupných čísel
- Replikovatelný
  - lze snadno získat tutéž posloupnost
  - Hodí se pro debugging a testování
- Nezávislé hodnoty
  - Posloupnost projde testem statistické nezávislosti
- Odpovídá požadovanému rozdělení
  - Střední hodnota, odchylka, histogram



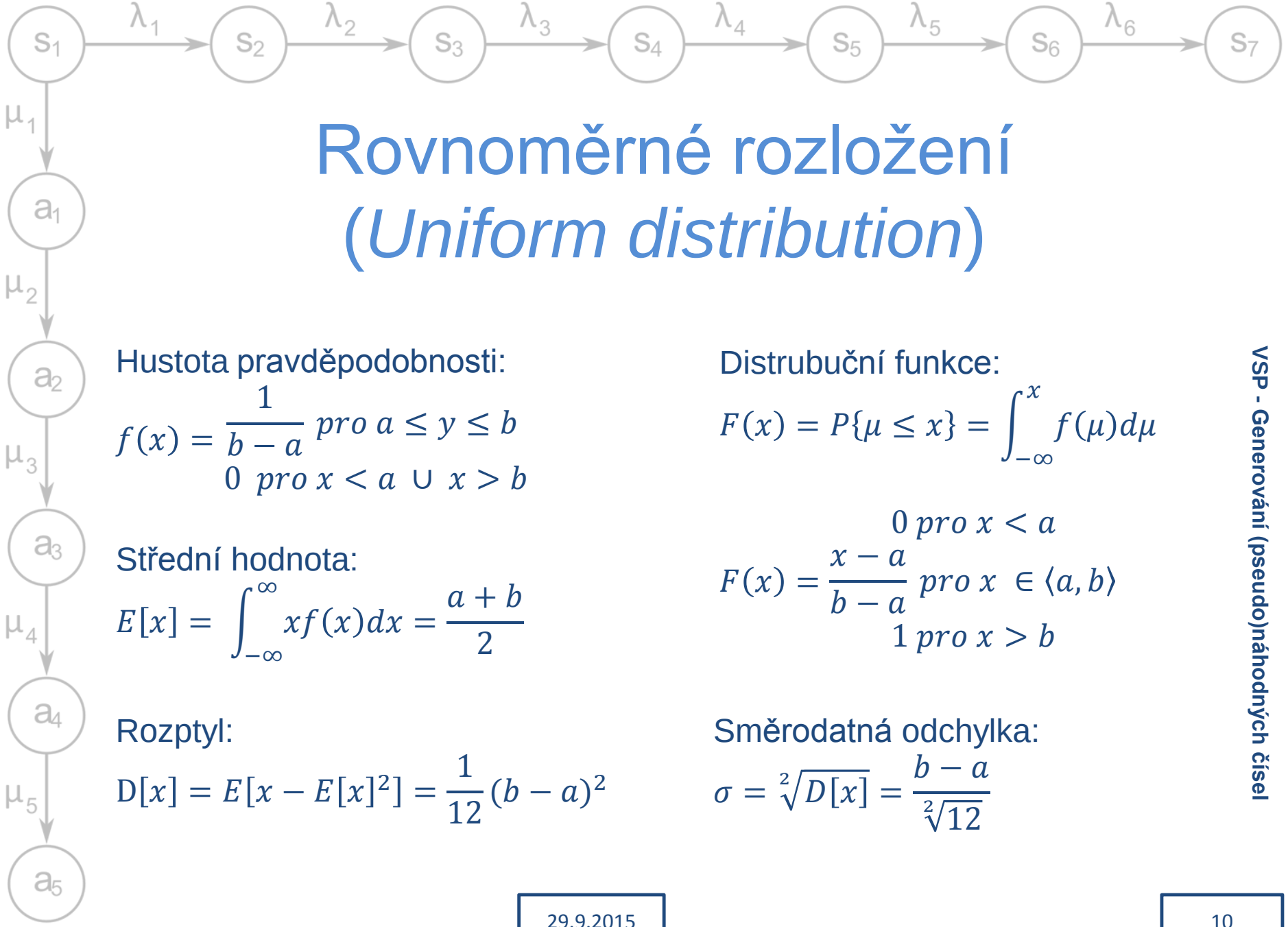
Rovnoměrné rozložení (*Uniform distribution*)

- Základ pro ostatní generátory
  - Parametry: minimum a maximum
  - Hodnoty ze zadaného rozsahu se stejnou pravděpodobností
  - Obvykle v rozsahu $\langle 0, 1 \rangle$ - normalizovaná podoba (neplést s normálním rozdělením!)

Hustota pravděpodobnosti

Distribuční funkce

Grafy:https://en.wikipedia.org/wiki/Uniform_distribution_(continuous)



Rovnoměrné rozložení (*Uniform distribution*)

Hustota pravděpodobnosti:

$$f(x) = \begin{matrix} \frac{1}{b-a}\ pro\ a \le y \le b \\ 0\ \ pro\ x < a\ \cup\ x > b \end{matrix}$$

Střední hodnota:

$$E[x] = \int_{-\infty}^{\infty} x f(x) dx = \frac{a+b}{2}$$

Rozptyl:

$$\mathrm{D}[x] = E[x - E[x]^2] = \frac{1}{12}(b-a)^2$$

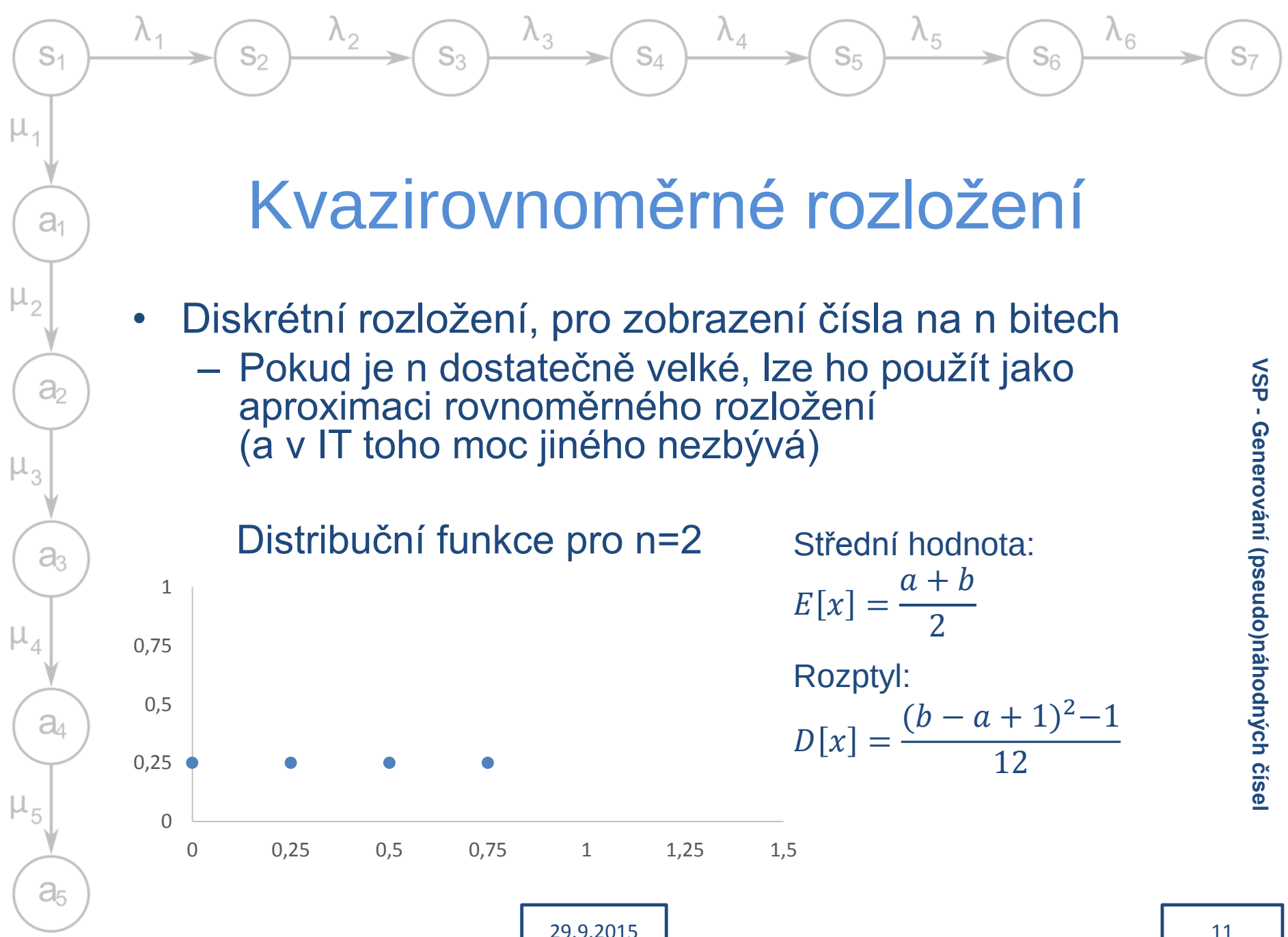
Distrubuční funkce:

$$F(x) = P\{\mu \le x\} = \int_{-\infty}^{x} f(\mu) d\mu$$

$$F(x) = \begin{matrix} 0\ pro\ x < a \\ \frac{x-a}{b-a}\ pro\ x\ \in \langle a, b \rangle \\ 1\ pro\ x > b \end{matrix}$$

Směrodatná odchylka:

$$\sigma = \sqrt[2]{D[x]} = \frac{b-a}{\sqrt[2]{12}}$$



Kvazirovnoměrné rozložení

- Diskrétní rozložení, pro zobrazení čísla na n bitech
  - Pokud je n dostatečně velké, lze ho použít jako aproximaci rovnoměrného rozložení (a v IT toho moc jiného nezbývá)

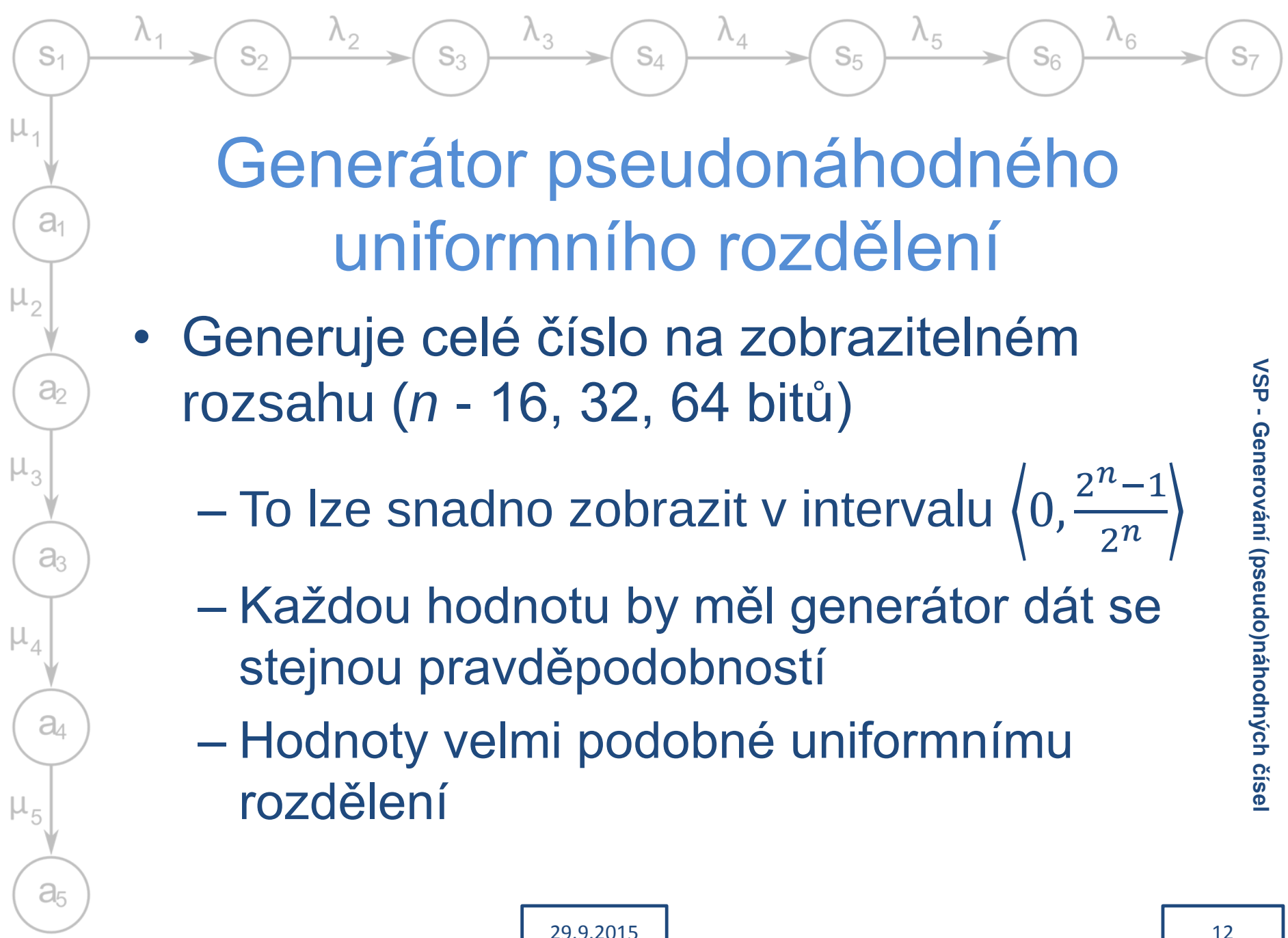
Distribuční funkce pro n=2

Střední hodnota:

$$E[x] = \frac{a+b}{2}$$

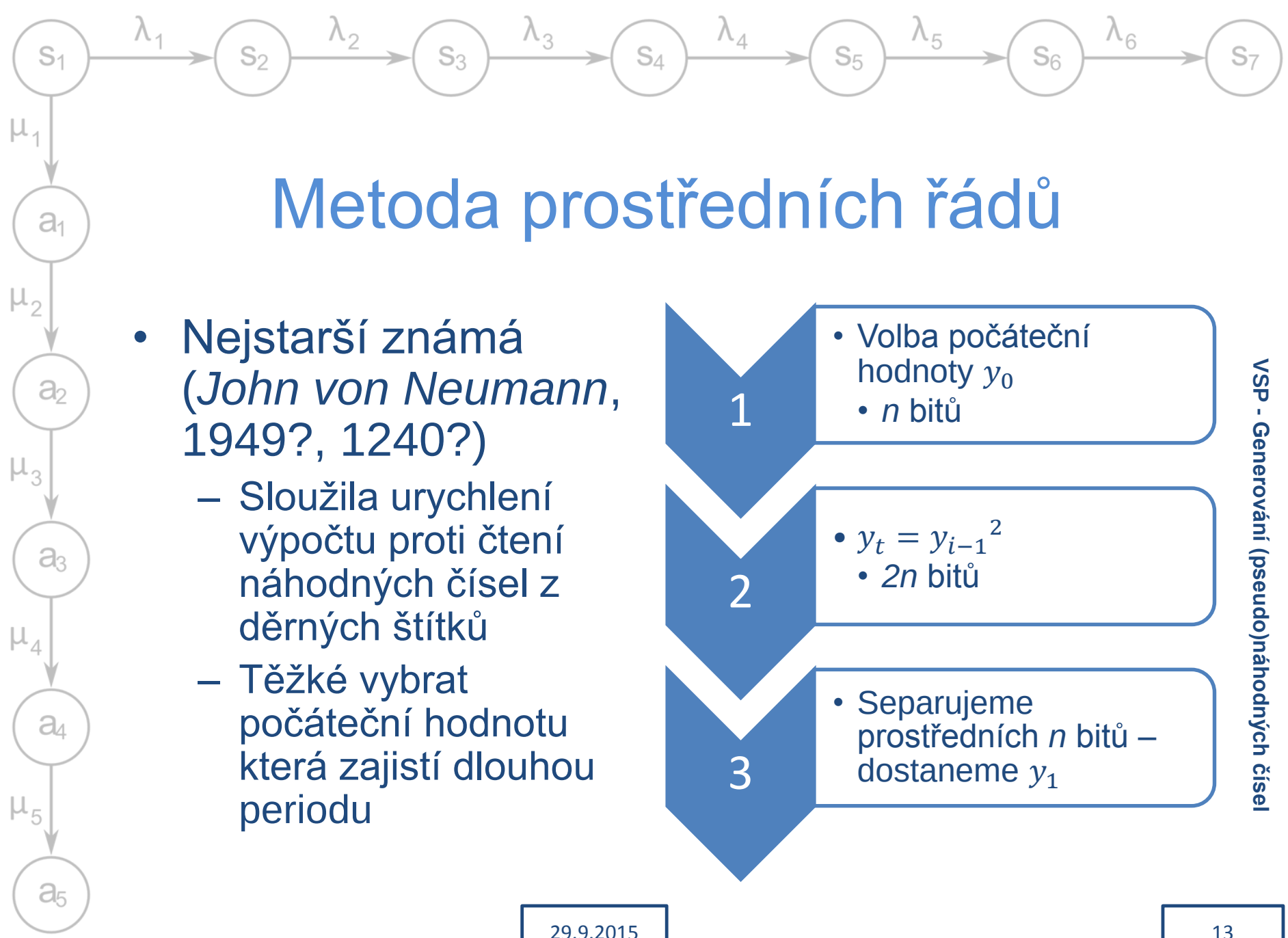
Rozptyl:

$$D[x] = \frac{(b-a+1)^2-1}{12}$$



Generátor pseudonáhodného uniformního rozdělení

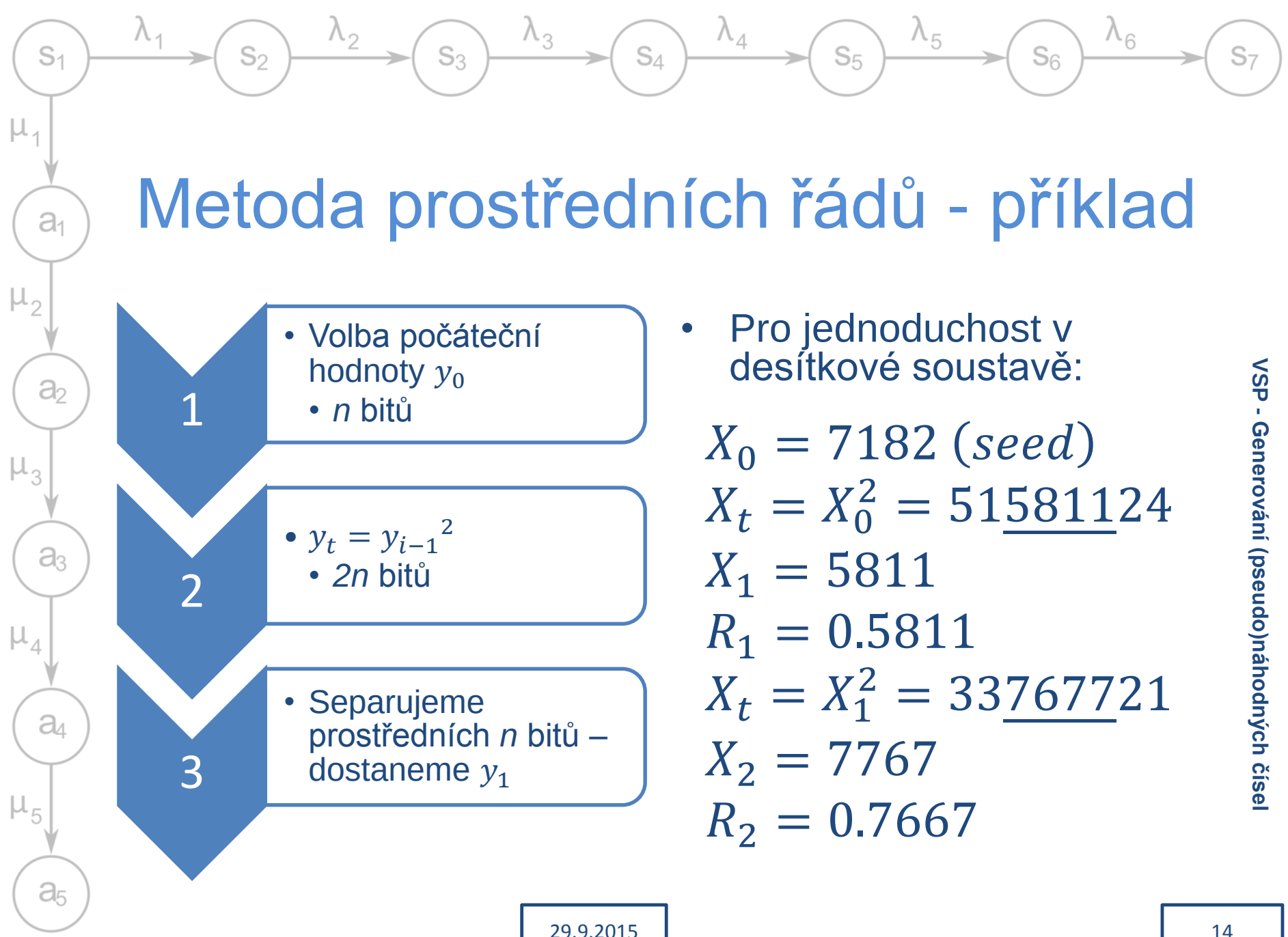
- Generuje celé číslo na zobrazitelném rozsahu (*n* - 16, 32, 64 bitů)
  - To lze snadno zobrazit v intervalu $\left\langle 0, \frac{2^n - 1}{2^n} \right\rangle$
  - Každou hodnotu by měl generátor dát se stejnou pravděpodobností
  - Hodnoty velmi podobné uniformnímu rozdělení



Metoda prostředních řádů

- Nejstarší známá (*John von Neumann*, 1949?, 1240?)
  - Sloužila urychlení výpočtu proti čtení náhodných čísel z děrných štítků
  - Těžké vybrat počáteční hodnotu která zajistí dlouhou periodu

1. • Volba počáteční hodnoty y_0
   • *n* bitů
2. • $y_t = {y_{i-1}}^2$
   • *2n* bitů
3. • Separujeme prostředních *n* bitů – dostaneme y_1

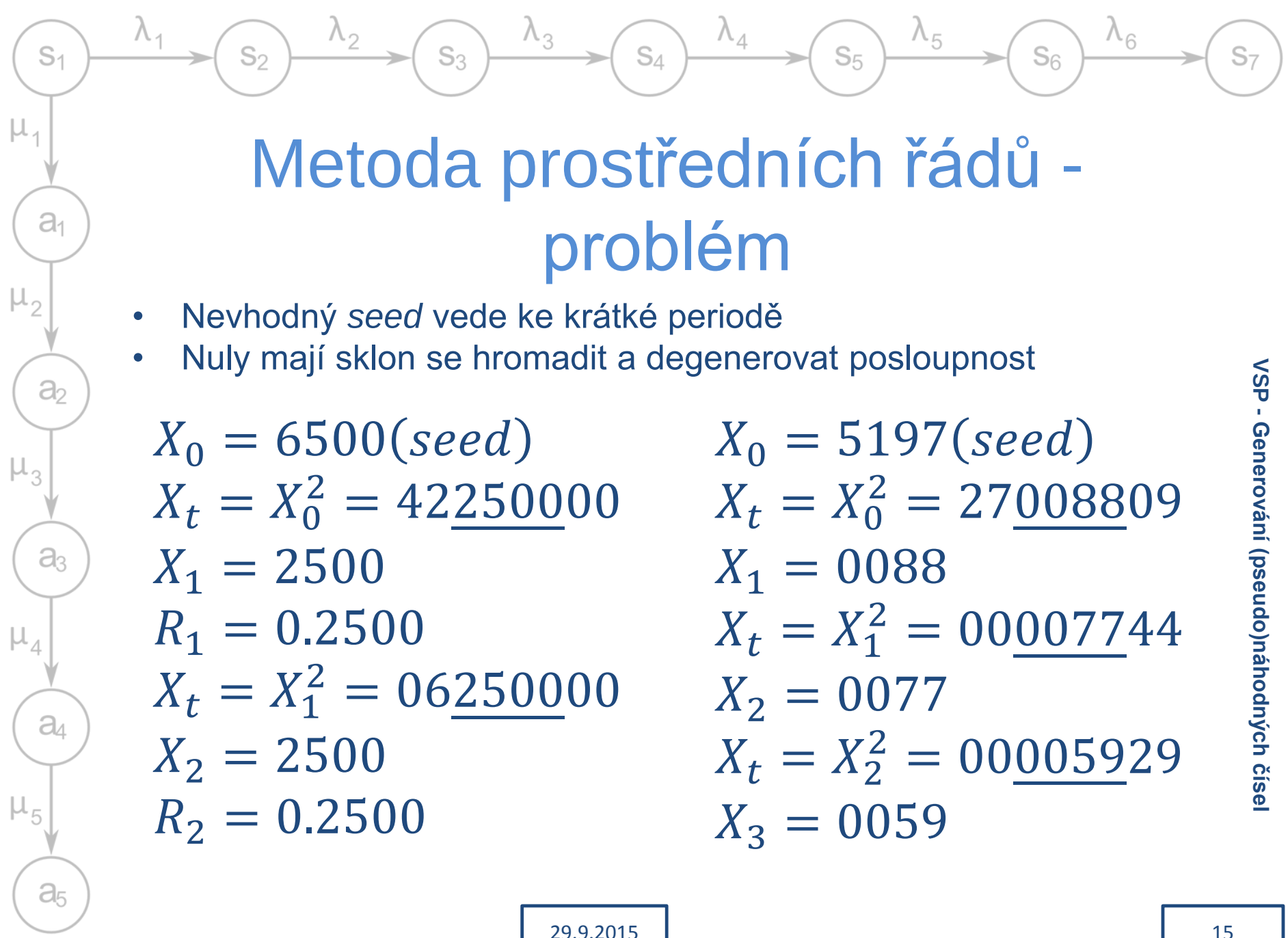


Metoda prostředních řádů - příklad

1. Volba počáteční hodnoty y_0
   - *n* bitů
2. $y_t = {y_{i-1}}^2$
   - *2n* bitů
3. Separujeme prostředních *n* bitů – dostaneme y_1

- Pro jednoduchost v desítkové soustavě:

$$X_0 = 7182\ (seed)$$
$$X_t = X_0^2 = 51\underline{5811}24$$
$$X_1 = 5811$$
$$R_1 = 0.5811$$
$$X_t = X_1^2 = 33\underline{7677}21$$
$$X_2 = 7767$$
$$R_2 = 0.7667$$

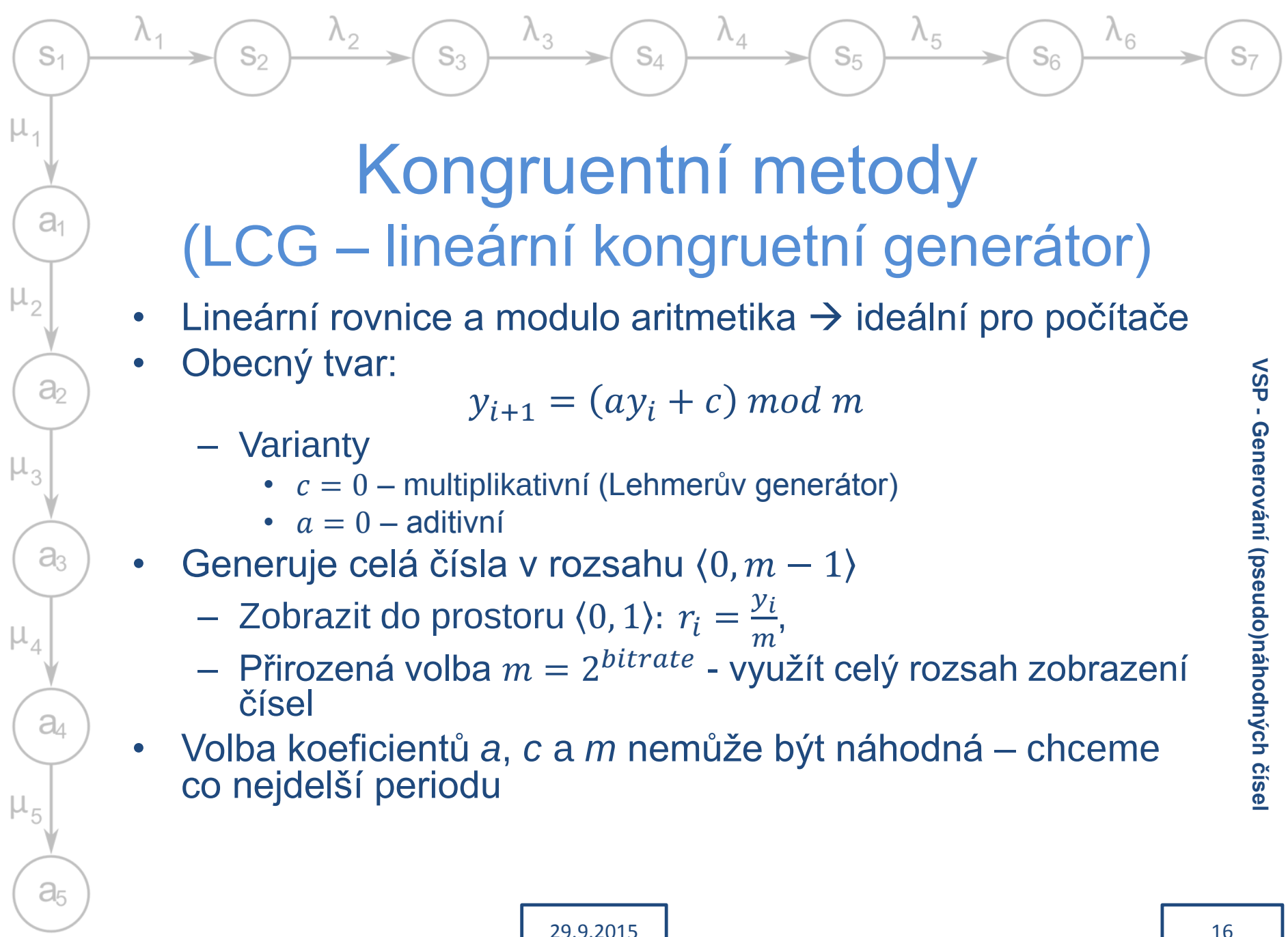


Metoda prostředních řádů - problém

- Nevhodný *seed* vede ke krátké periodě
- Nuly mají sklon se hromadit a degenerovat posloupnost

$X_0 = 6500(seed)$
$X_t = X_0^2 = 42\underline{2500}00$
$X_1 = 2500$
$R_1 = 0.2500$
$X_t = X_1^2 = 06\underline{2500}00$
$X_2 = 2500$
$R_2 = 0.2500$

$X_0 = 5197(seed)$
$X_t = X_0^2 = 27\underline{0088}09$
$X_1 = 0088$
$X_t = X_1^2 = 00\underline{0077}44$
$X_2 = 0077$
$X_t = X_2^2 = 00\underline{0059}29$
$X_3 = 0059$

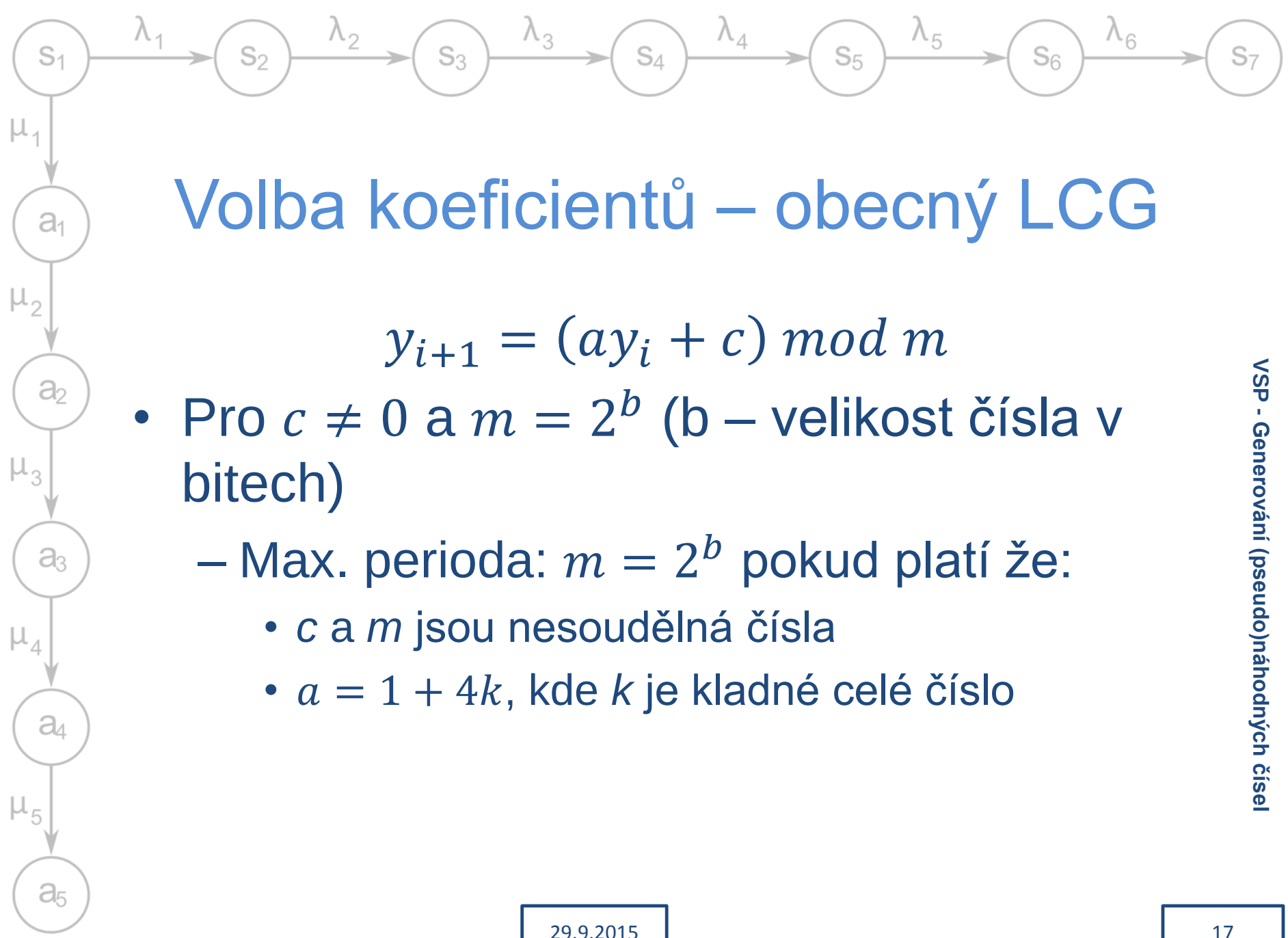


Kongruentní metody
(LCG – lineární kongruetní generátor)

- Lineární rovnice a modulo aritmetika → ideální pro počítače
- Obecný tvar:

$$y_{i+1} = (ay_i + c)\ mod\ m$$

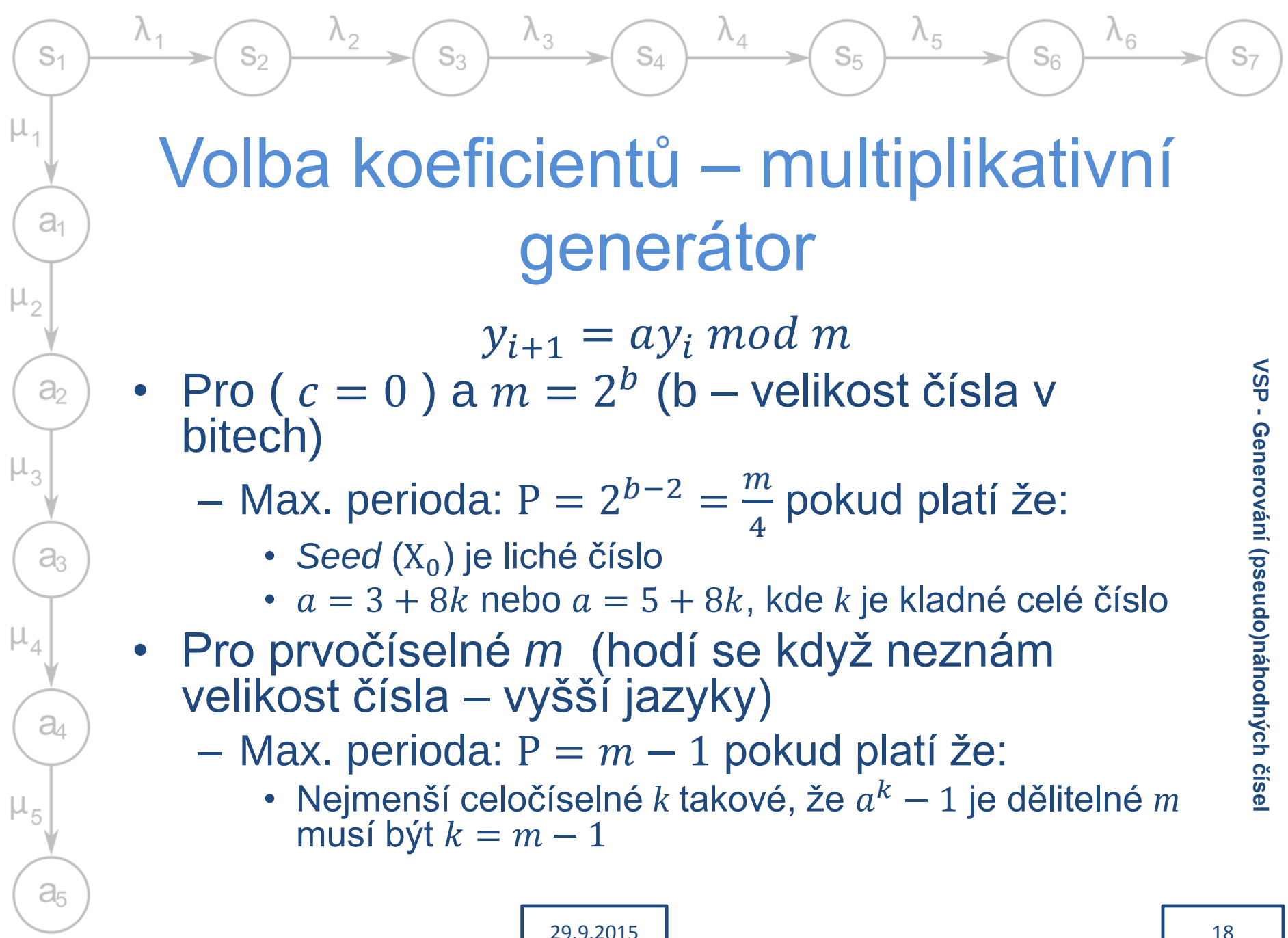
  - Varianty
    - $c = 0$ – multiplikativní (Lehmerův generátor)
    - $a = 0$ – aditivní
- Generuje celá čísla v rozsahu $\langle 0, m-1 \rangle$
  - Zobrazit do prostoru $\langle 0, 1 \rangle$: $r_i = \frac{y_i}{m}$,
  - Přirozená volba $m = 2^{bitrate}$ - využít celý rozsah zobrazení čísel
- Volba koeficientů *a*, *c* a *m* nemůže být náhodná – chceme co nejdelší periodu



Volba koeficientů – obecný LCG

$$y_{i+1} = (ay_i + c)\ mod\ m$$

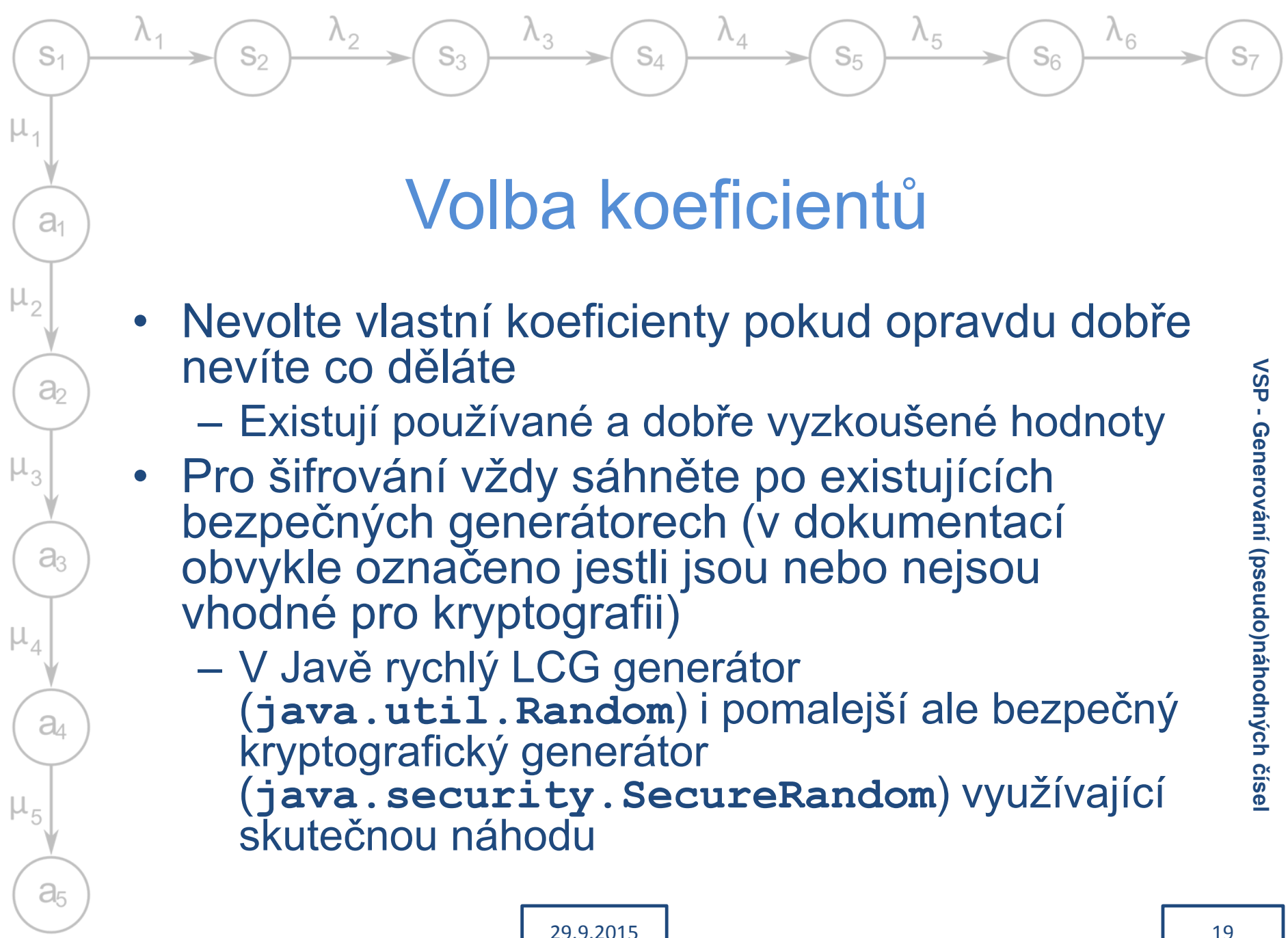
- Pro $c \neq 0$ a $m = 2^b$ (b – velikost čísla v bitech)
  - Max. perioda: $m = 2^b$ pokud platí že:
    - *c* a *m* jsou nesoudělná čísla
    - $a = 1 + 4k$, kde *k* je kladné celé číslo



Volba koeficientů – multiplikativní generátor

$$y_{i+1} = a y_i \; mod \; m$$

- Pro ($c = 0$) a $m = 2^b$ (b – velikost čísla v bitech)
  - Max. perioda: $\mathrm{P} = 2^{b-2} = \frac{m}{4}$ pokud platí že:
    - *Seed* (X_0) je liché číslo
    - $a = 3 + 8k$ nebo $a = 5 + 8k$, kde k je kladné celé číslo
- Pro prvočíselné m (hodí se když neznám velikost čísla – vyšší jazyky)
  - Max. perioda: $\mathrm{P} = m - 1$ pokud platí že:
    - Nejmenší celočíselné k takové, že $a^k - 1$ je dělitelné m musí být $k = m - 1$

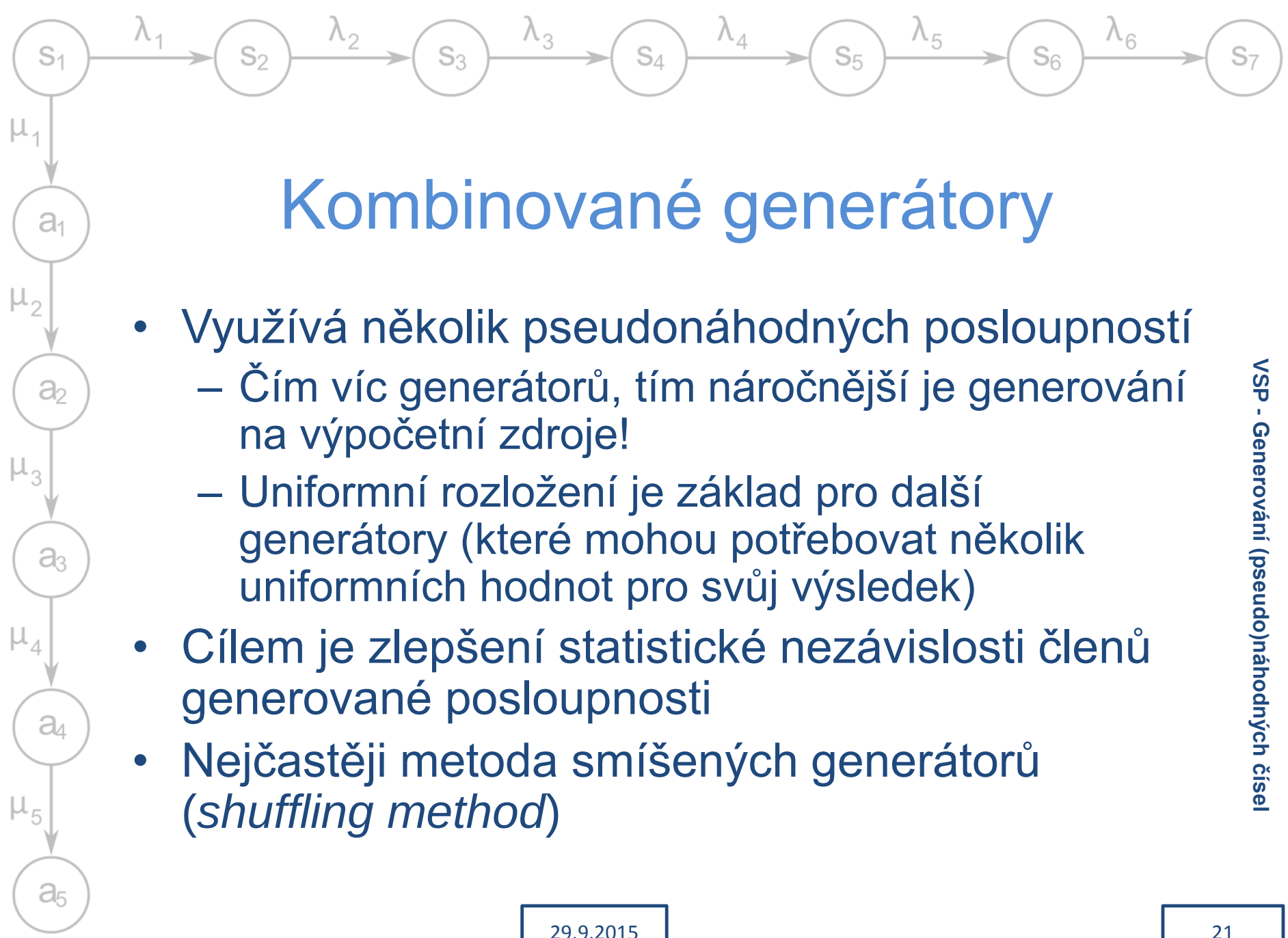


Volba koeficientů

- Nevolte vlastní koeficienty pokud opravdu dobře nevíte co děláte
  - Existují používané a dobře vyzkoušené hodnoty
- Pro šifrování vždy sáhněte po existujících bezpečných generátorech (v dokumentací obvykle označeno jestli jsou nebo nejsou vhodné pro kryptografii)
  - V Javě rychlý LCG generátor (**`java.util.Random`**) i pomalejší ale bezpečný kryptografický generátor (**`java.security.SecureRandom`**) využívající skutečnou náhodu

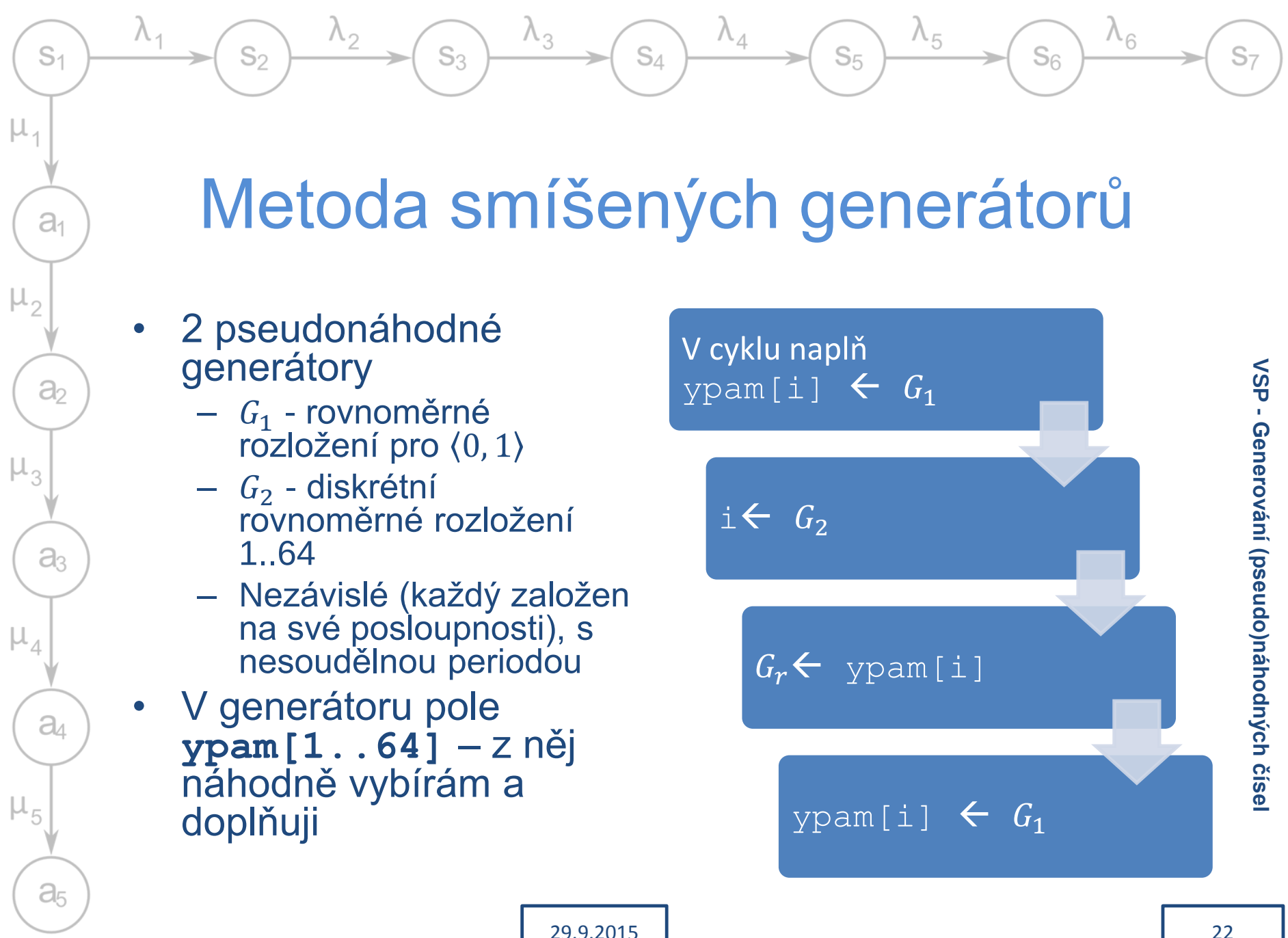
Příklady koeficientů pro LCG

Zdroj	*m*	*a*	*c*	Použité bity
GCC glibc	2^{31}	1103515245	12345	30..0
ISO/IEC 9899 (C 2011)	2^{32}	1103515245	12345	30..16
`java.util.Random`	2^{48}	25214903917	11	47...16
Apple CarbonLib	2^{31}-1	16807	0	
C Newlib, MUSL	2^{64}	6364136223846793005	1	63...32
Visual Basic	2^{24}	1140671485	12820163	



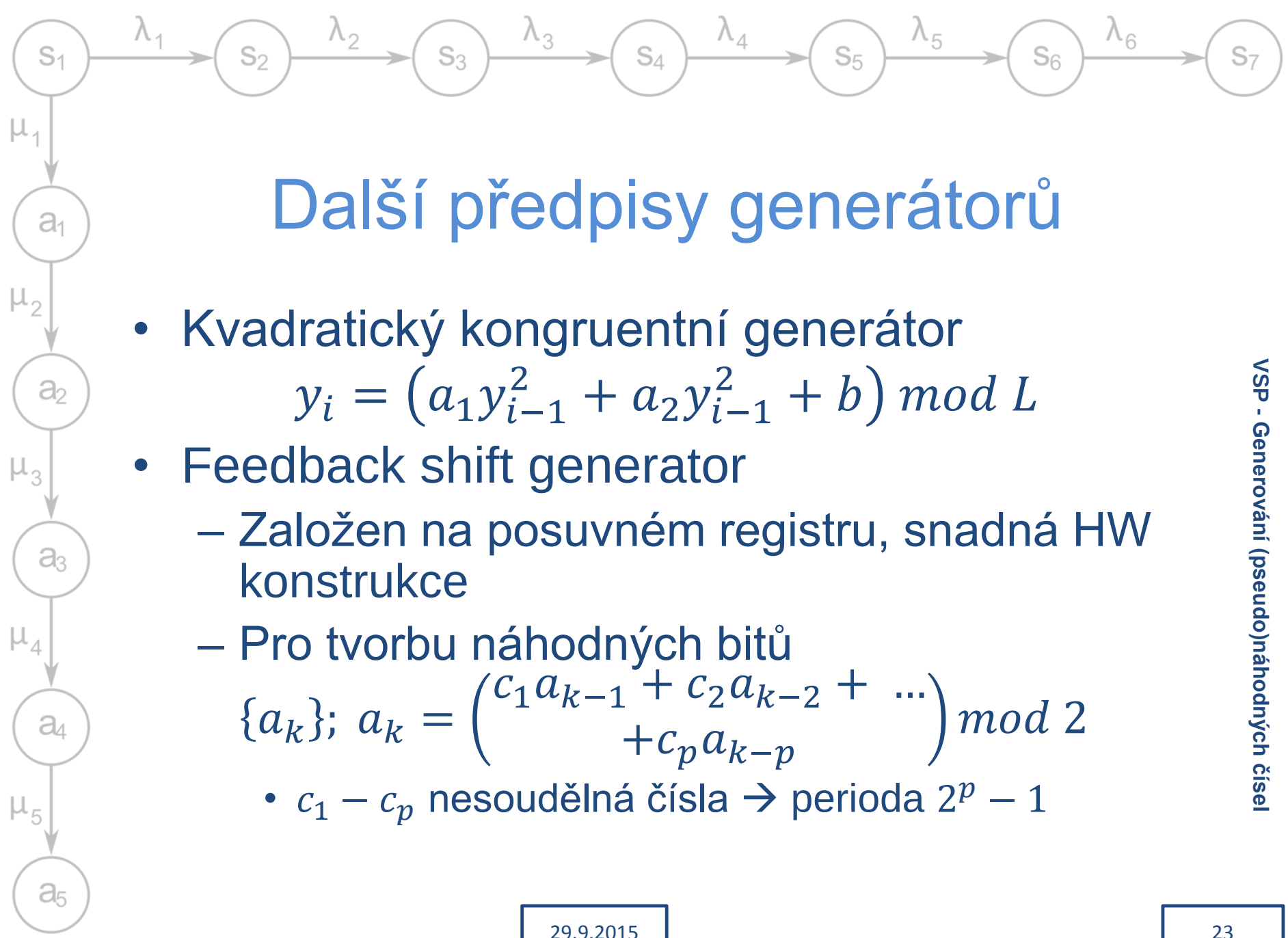
Kombinované generátory

- Využívá několik pseudonáhodných posloupností
  - Čím víc generátorů, tím náročnější je generování na výpočetní zdroje!
  - Uniformní rozložení je základ pro další generátory (které mohou potřebovat několik uniformních hodnot pro svůj výsledek)
- Cílem je zlepšení statistické nezávislosti členů generované posloupnosti
- Nejčastěji metoda smíšených generátorů (*shuffling method*)



Metoda smíšených generátorů

- 2 pseudonáhodné generátory
  - G_1 - rovnoměrné rozložení pro $\langle 0, 1 \rangle$
  - G_2 - diskrétní rovnoměrné rozložení 1..64
  - Nezávislé (každý založen na své posloupnosti), s nesoudělnou periodou
- V generátoru pole **`ypam[1..64]`** – z něj náhodně vybírám a doplňuji



Další předpisy generátorů

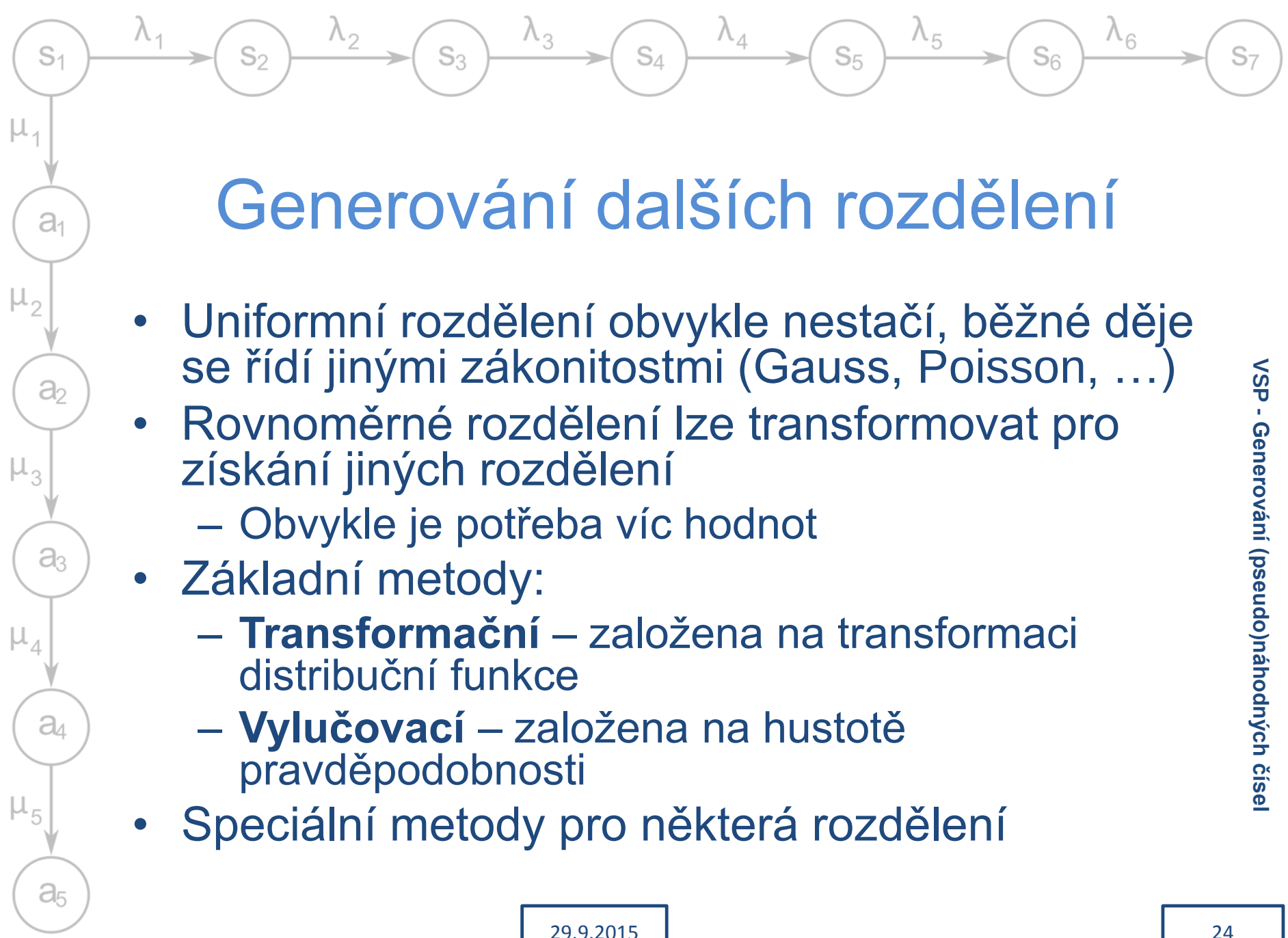
- Kvadratický kongruentní generátor

$$y_i = \left(a_1 y_{i-1}^2 + a_2 y_{i-1}^2 + b\right) mod\ L$$

- Feedback shift generator
  - Založen na posuvném registru, snadná HW konstrukce
  - Pro tvorbu náhodných bitů

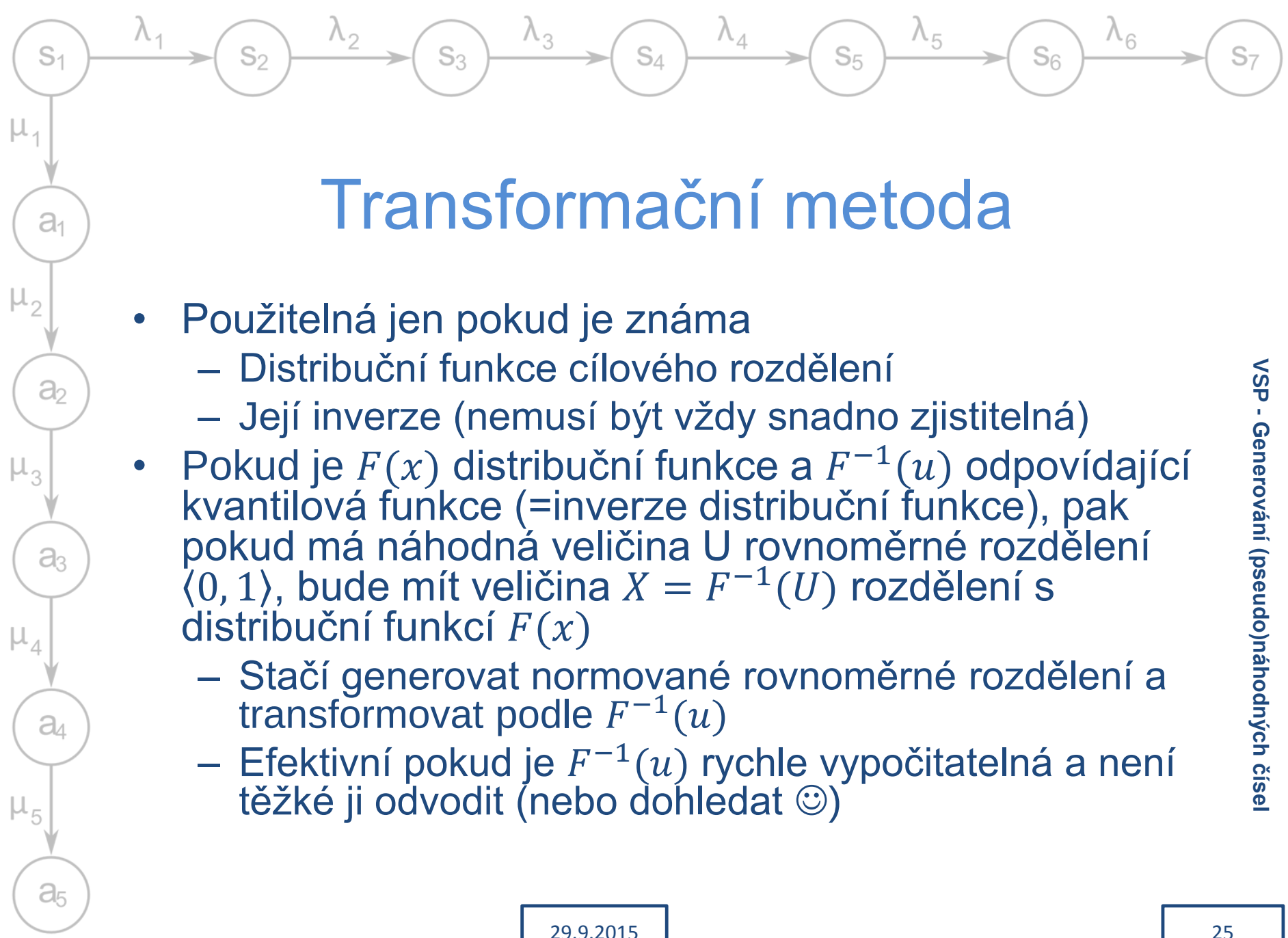
$$\{a_k\};\ a_k = \begin{pmatrix} c_1 a_{k-1} + c_2 a_{k-2} + \ \dots \\ + c_p a_{k-p} \end{pmatrix} mod\ 2$$

    - $c_1 - c_p$ nesoudělná čísla → perioda $2^p - 1$



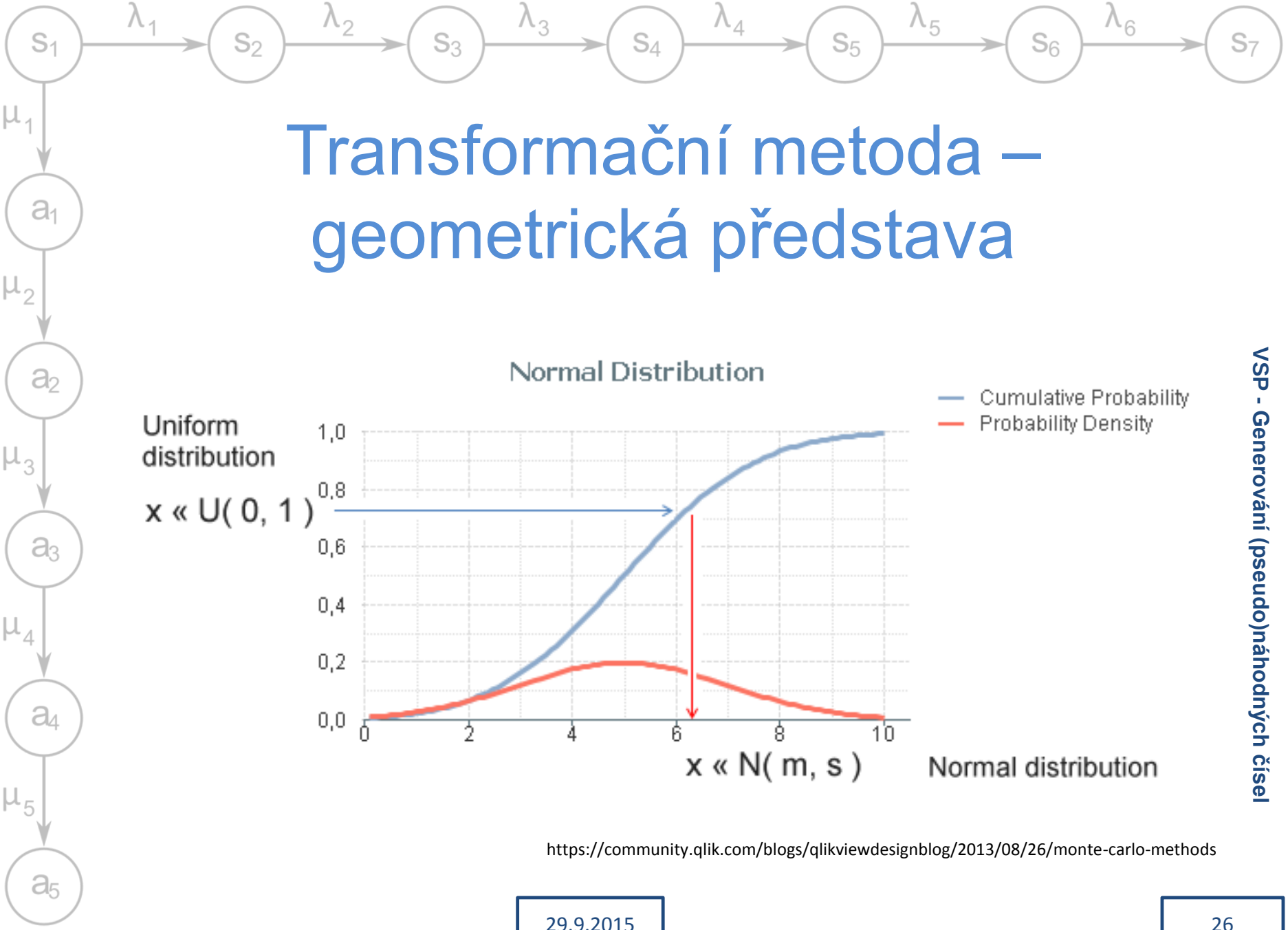
Generování dalších rozdělení

- Uniformní rozdělení obvykle nestačí, běžné děje se řídí jinými zákonitostmi (Gauss, Poisson, …)
- Rovnoměrné rozdělení lze transformovat pro získání jiných rozdělení
  - Obvykle je potřeba víc hodnot
- Základní metody:
  - **Transformační** – založena na transformaci distribuční funkce
  - **Vylučovací** – založena na hustotě pravděpodobnosti
- Speciální metody pro některá rozdělení



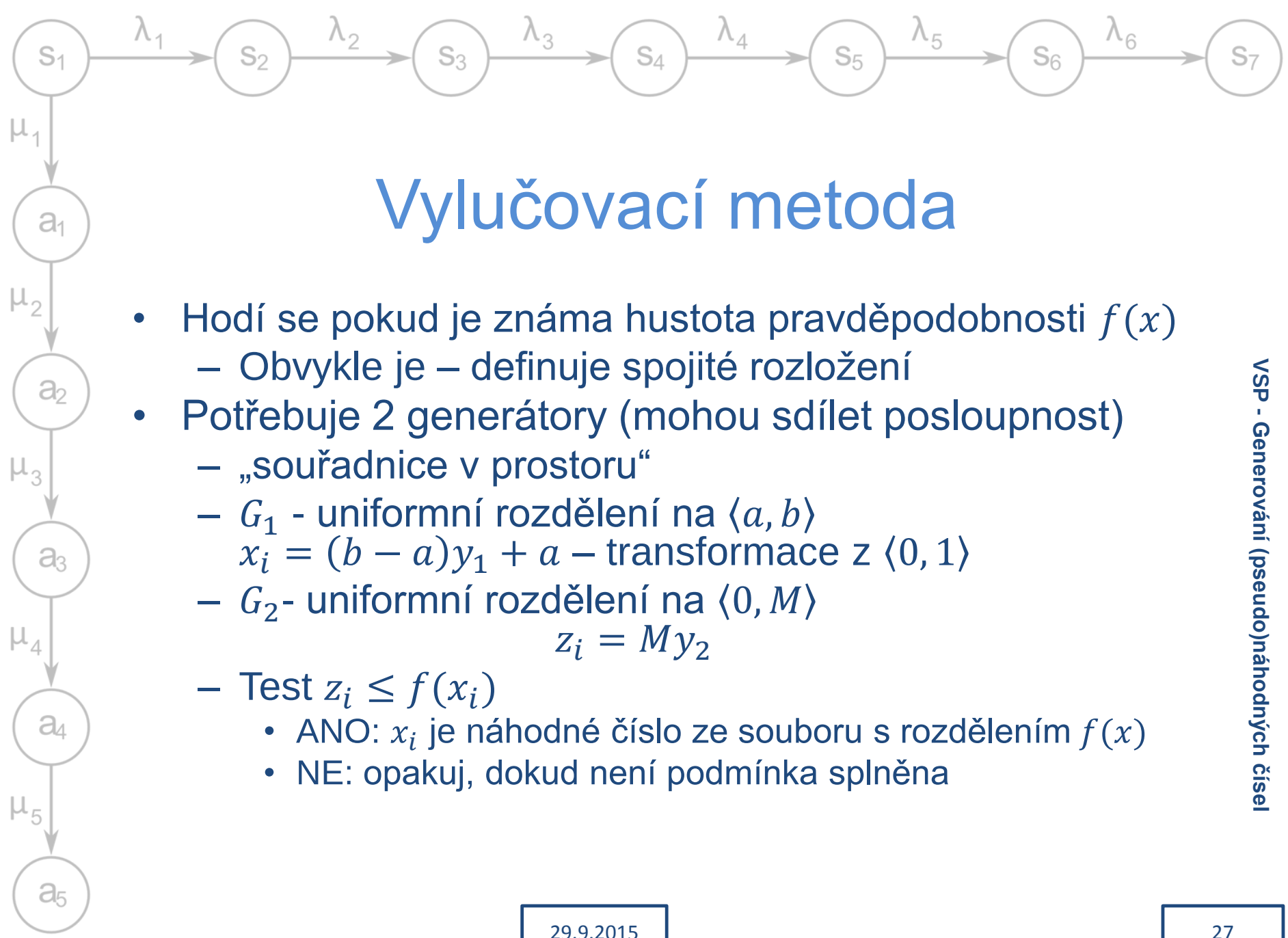
Transformační metoda

- Použitelná jen pokud je známa
  - Distribuční funkce cílového rozdělení
  - Její inverze (nemusí být vždy snadno zjistitelná)
- Pokud je $F(x)$ distribuční funkce a $F^{-1}(u)$ odpovídající kvantilová funkce (=inverze distribuční funkce), pak pokud má náhodná veličina U rovnoměrné rozdělení $\langle 0, 1 \rangle$, bude mít veličina $X = F^{-1}(U)$ rozdělení s distribuční funkcí $F(x)$
  - Stačí generovat normované rovnoměrné rozdělení a transformovat podle $F^{-1}(u)$
  - Efektivní pokud je $F^{-1}(u)$ rychle vypočitatelná a není těžké ji odvodit (nebo dohledat ☺)



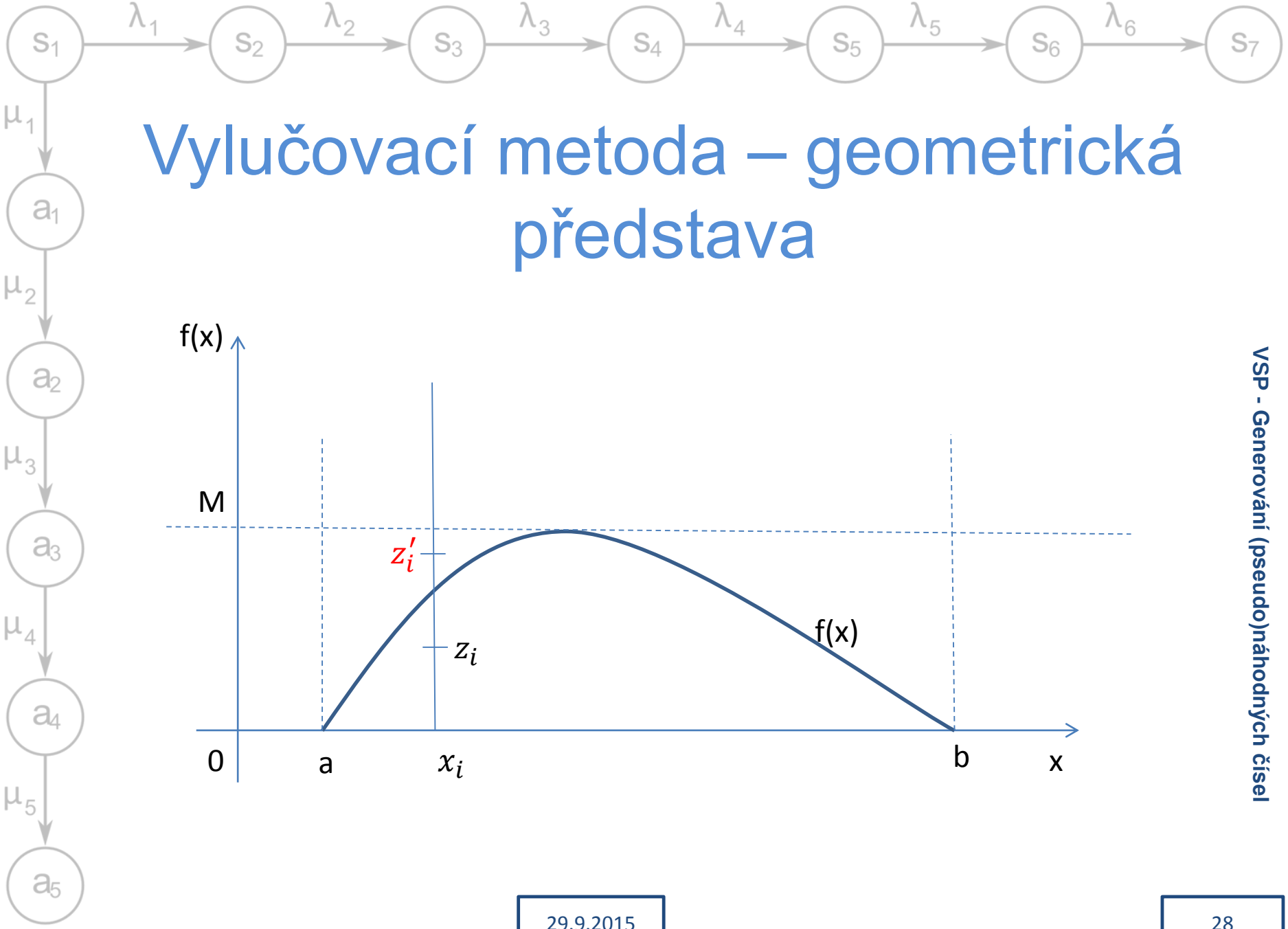
Transformační metoda – geometrická představa

https://community.qlik.com/blogs/qlikviewdesignblog/2013/08/26/monte-carlo-methods

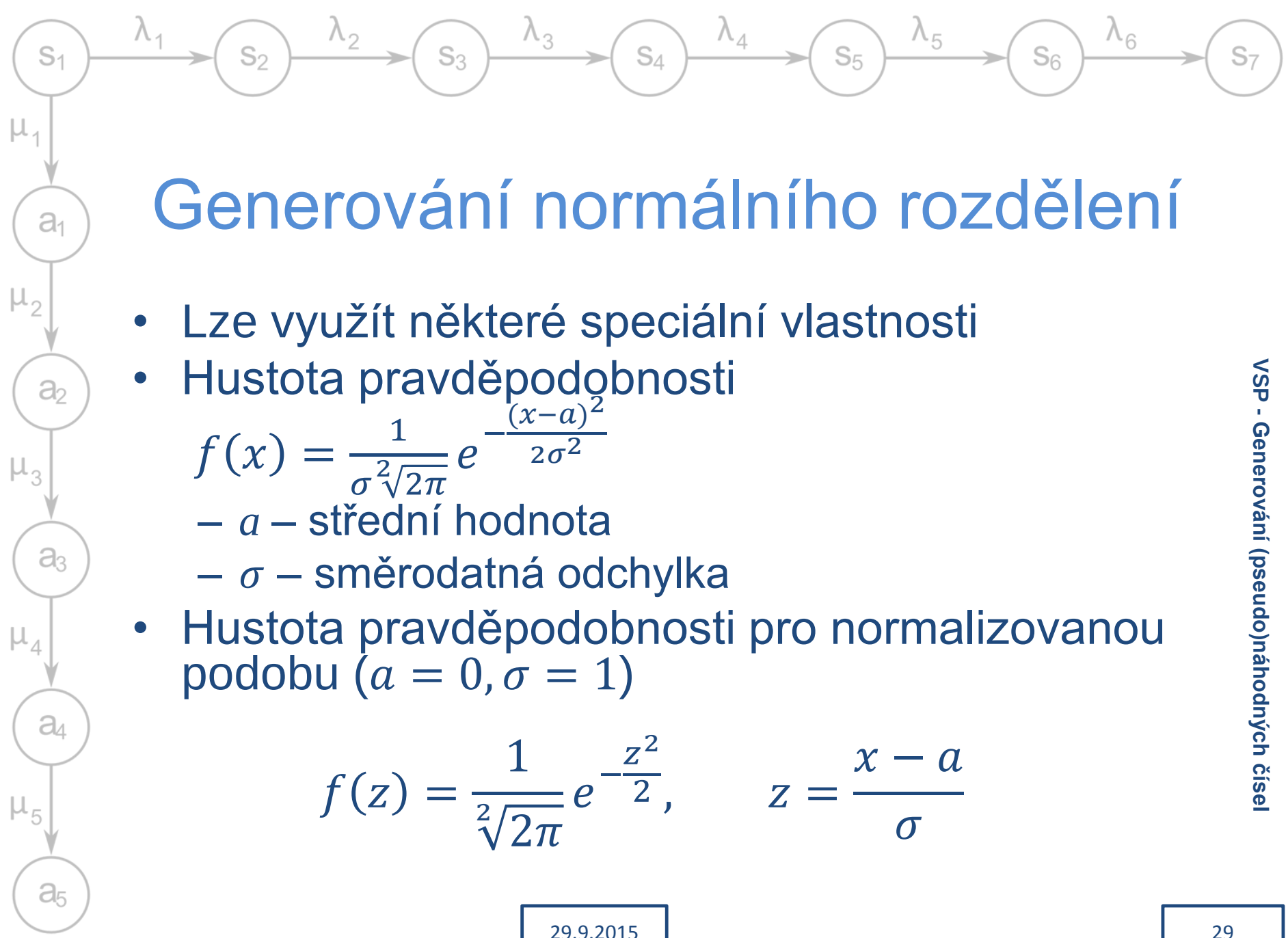


Vylučovací metoda

- Hodí se pokud je známa hustota pravděpodobnosti $f(x)$
  - Obvykle je – definuje spojité rozložení
- Potřebuje 2 generátory (mohou sdílet posloupnost)
  - „souřadnice v prostoru“
  - G_1 - uniformní rozdělení na $\langle a, b \rangle$
    $x_i = (b - a)y_1 + a$ – transformace z $\langle 0, 1 \rangle$
  - G_2- uniformní rozdělení na $\langle 0, M \rangle$
    $$z_i = My_2$$
  - Test $z_i \leq f(x_i)$
    - ANO: x_i je náhodné číslo ze souboru s rozdělením $f(x)$
    - NE: opakuj, dokud není podmínka splněna



Vylučovací metoda – geometrická představa



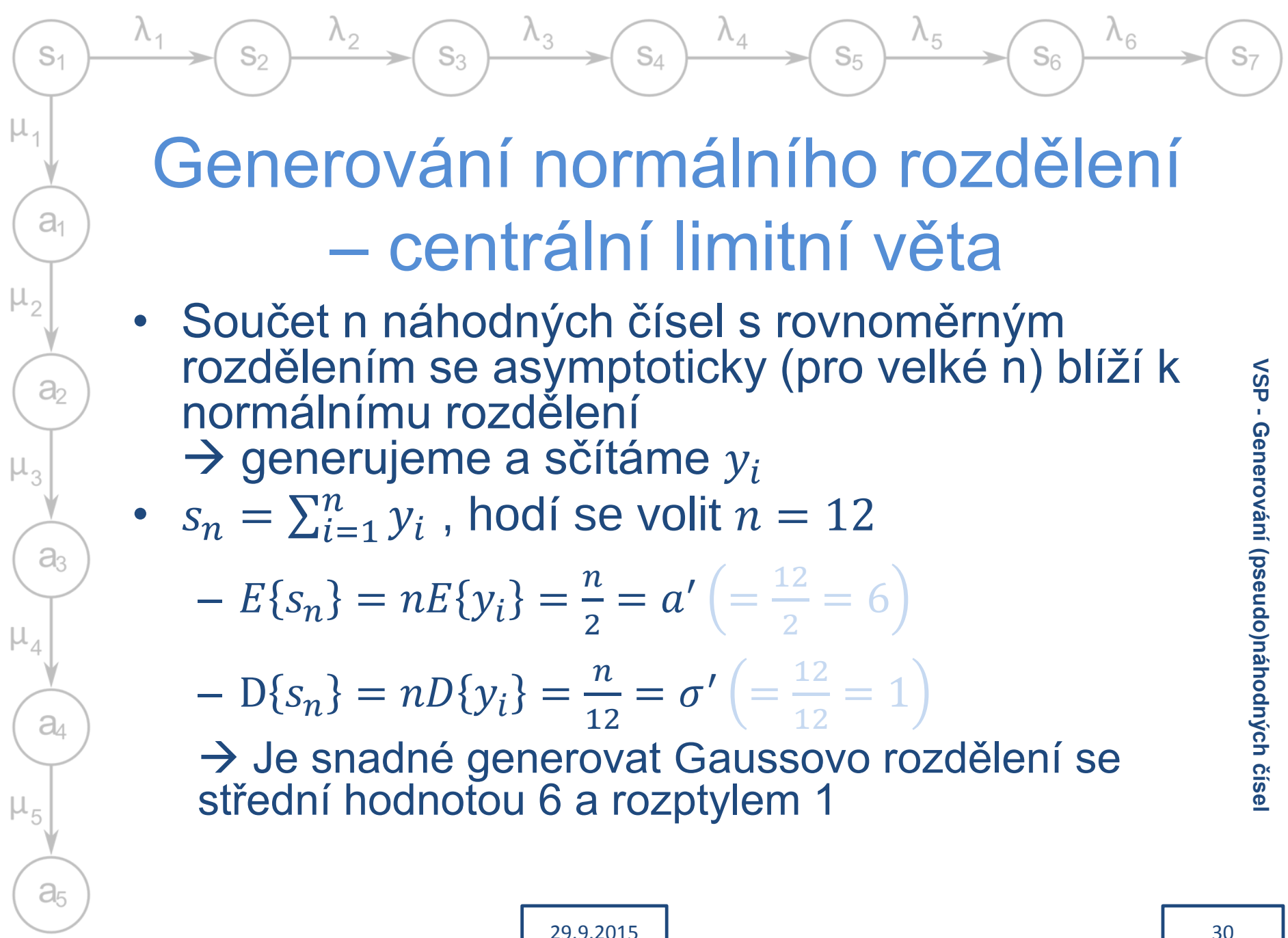
Generování normálního rozdělení

- Lze využít některé speciální vlastnosti
- Hustota pravděpodobnosti

$$f(x) = \frac{1}{\sigma\sqrt[2]{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

  - a – střední hodnota
  - σ – směrodatná odchylka
- Hustota pravděpodobnosti pro normalizovanou podobu ($a = 0, \sigma = 1$)

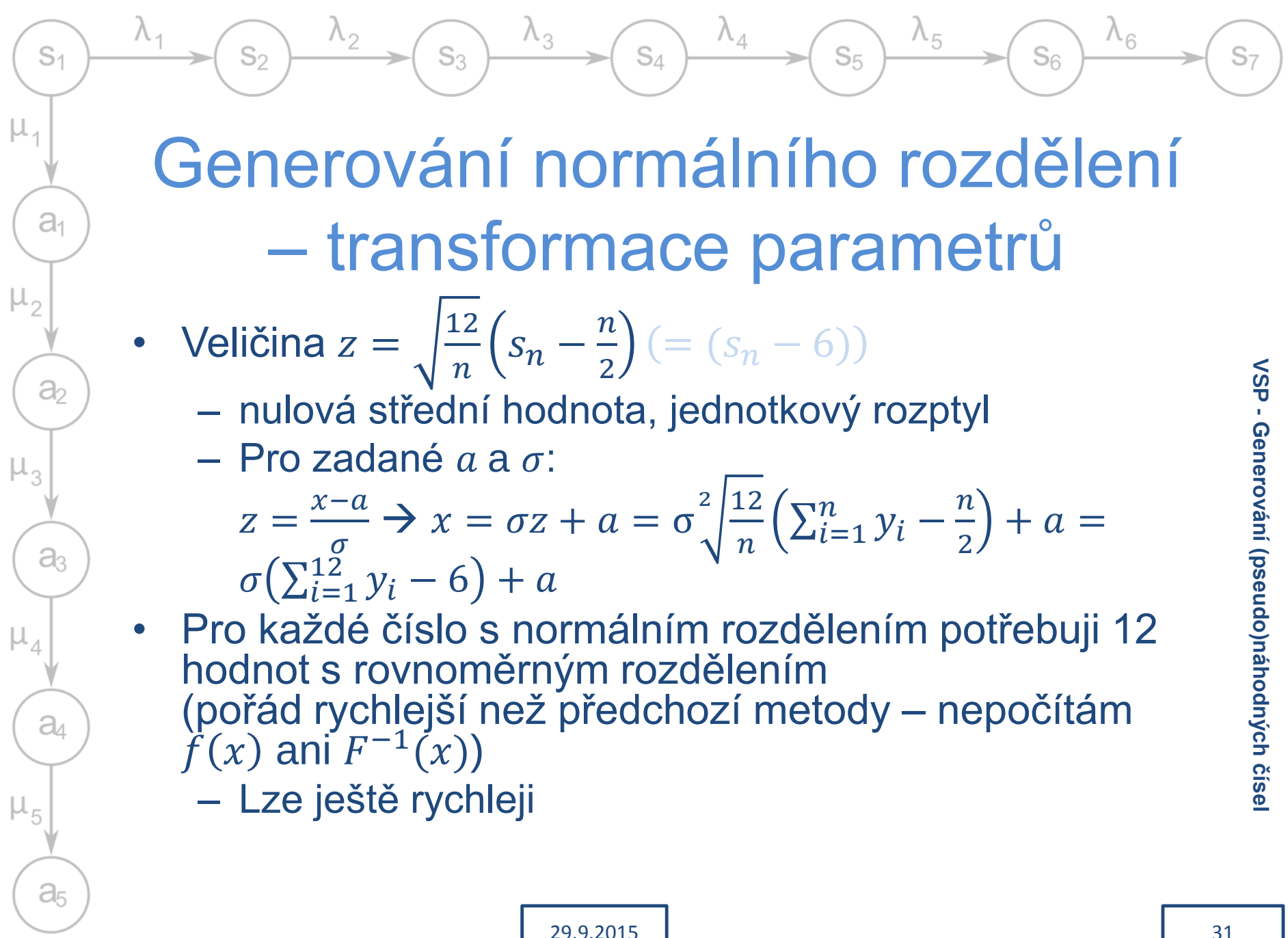
$$f(z) = \frac{1}{\sqrt[2]{2\pi}} e^{-\frac{z^2}{2}}, \qquad z = \frac{x-a}{\sigma}$$



Generování normálního rozdělení – centrální limitní věta

- Součet n náhodných čísel s rovnoměrným rozdělením se asymptoticky (pro velké n) blíží k normálnímu rozdělení
  → generujeme a sčítáme y_i
- $s_n = \sum_{i=1}^{n} y_i$, hodí se volit $n = 12$
  - $E\{s_n\} = nE\{y_i\} = \frac{n}{2} = a' \left(= \frac{12}{2} = 6\right)$
  - $\mathrm{D}\{s_n\} = nD\{y_i\} = \frac{n}{12} = \sigma' \left(= \frac{12}{12} = 1\right)$

  → Je snadné generovat Gaussovo rozdělení se střední hodnotou 6 a rozptylem 1

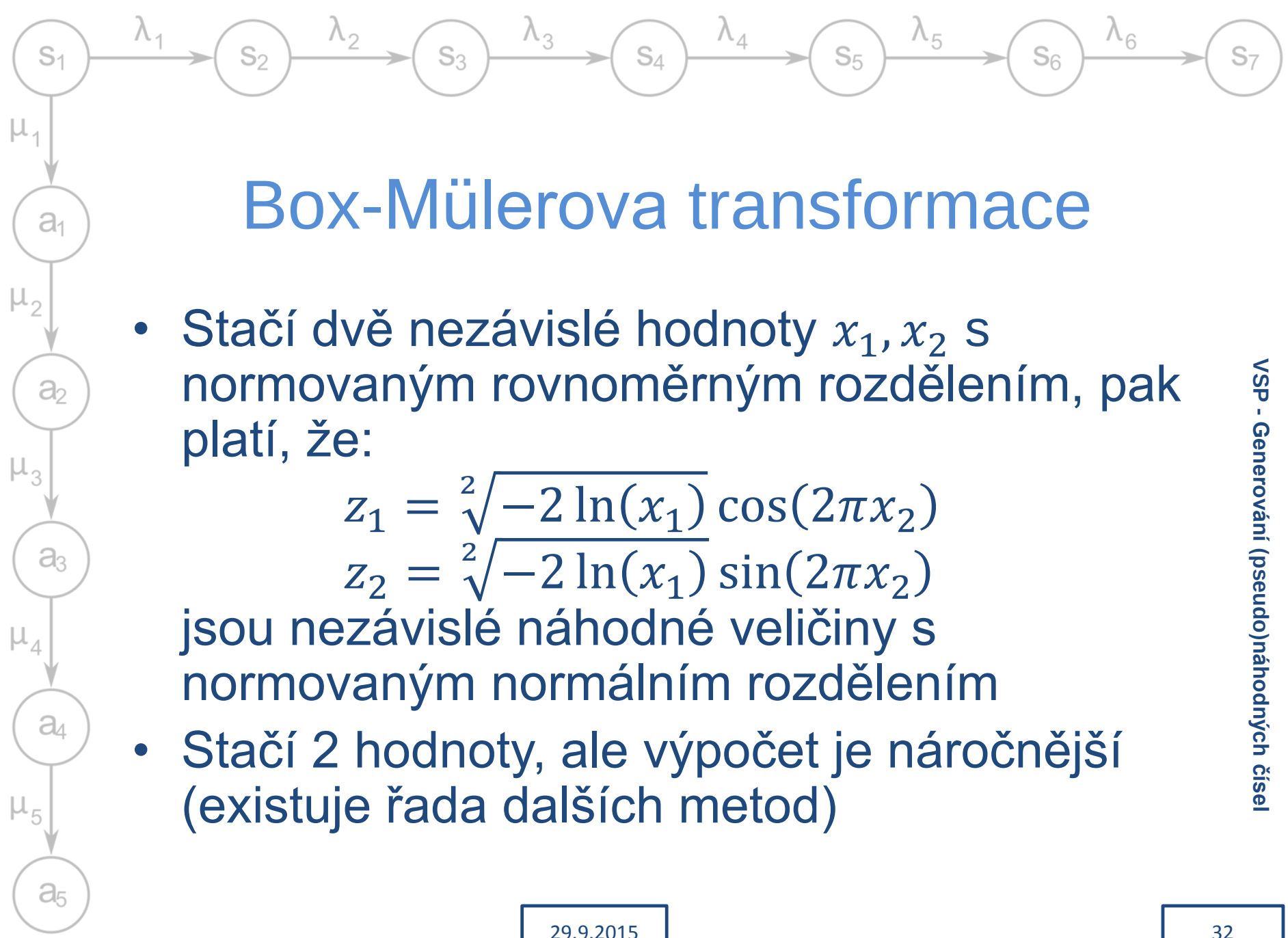


Generování normálního rozdělení – transformace parametrů

- Veličina $z = \sqrt{\frac{12}{n}}\left(s_n - \frac{n}{2}\right) \; (= (s_n - 6))$
  - nulová střední hodnota, jednotkový rozptyl
  - Pro zadané a a σ:

    $z = \frac{x-a}{\sigma} \rightarrow x = \sigma z + a = \sigma \sqrt{\frac{12}{n}}\left(\sum_{i=1}^{n} y_i - \frac{n}{2}\right) + a =$

    $\sigma\left(\sum_{i=1}^{12} y_i - 6\right) + a$
- Pro každé číslo s normálním rozdělením potřebuji 12 hodnot s rovnoměrným rozdělením (pořád rychlejší než předchozí metody – nepočítám $f(x)$ ani $F^{-1}(x)$)
  - Lze ještě rychleji

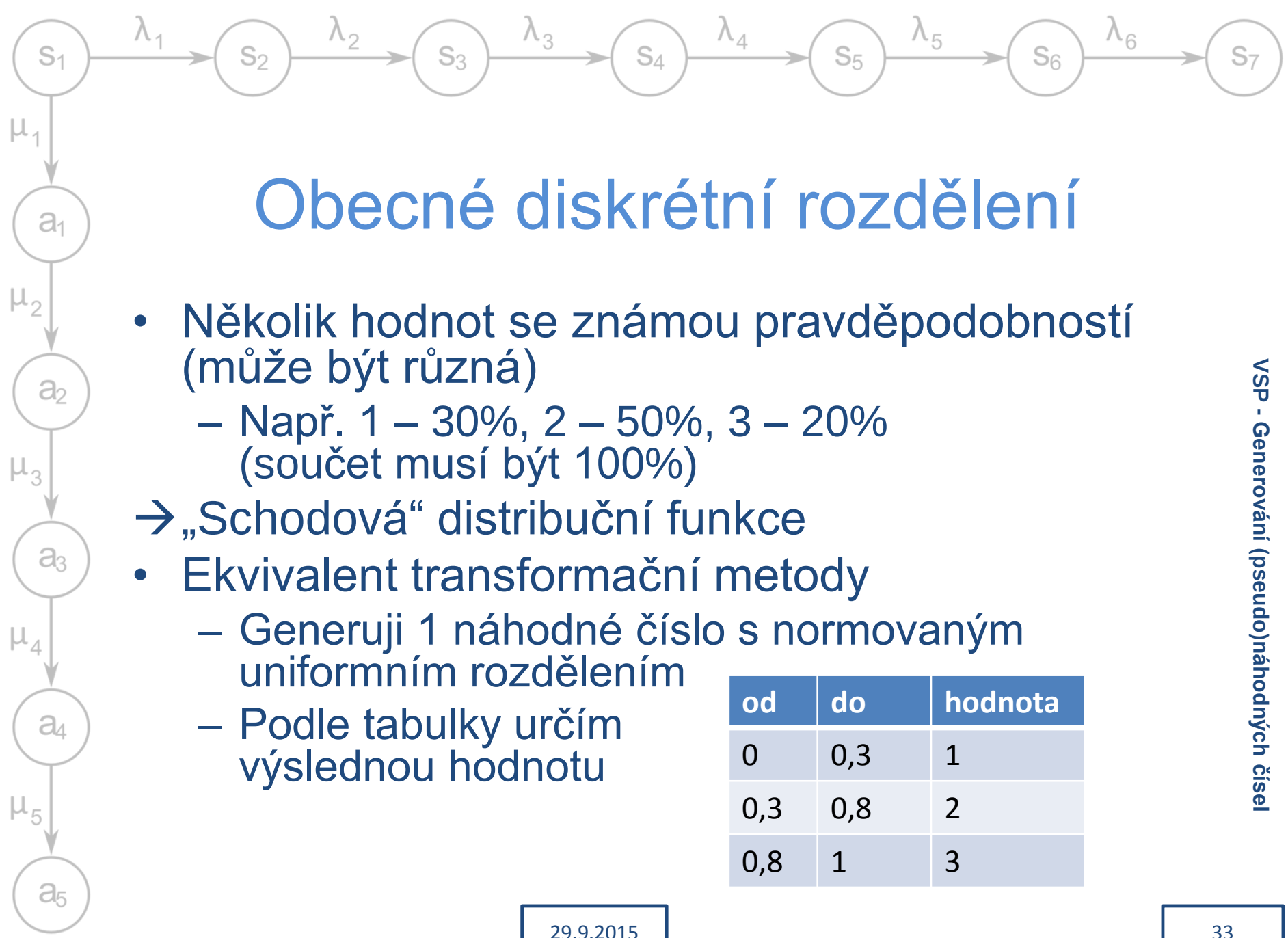


Box-Mülerova transformace

- Stačí dvě nezávislé hodnoty x_1, x_2 s normovaným rovnoměrným rozdělením, pak platí, že:

$$z_1 = \sqrt[2]{-2\ln(x_1)}\cos(2\pi x_2)$$
$$z_2 = \sqrt[2]{-2\ln(x_1)}\sin(2\pi x_2)$$

  jsou nezávislé náhodné veličiny s normovaným normálním rozdělením
- Stačí 2 hodnoty, ale výpočet je náročnější (existuje řada dalších metod)



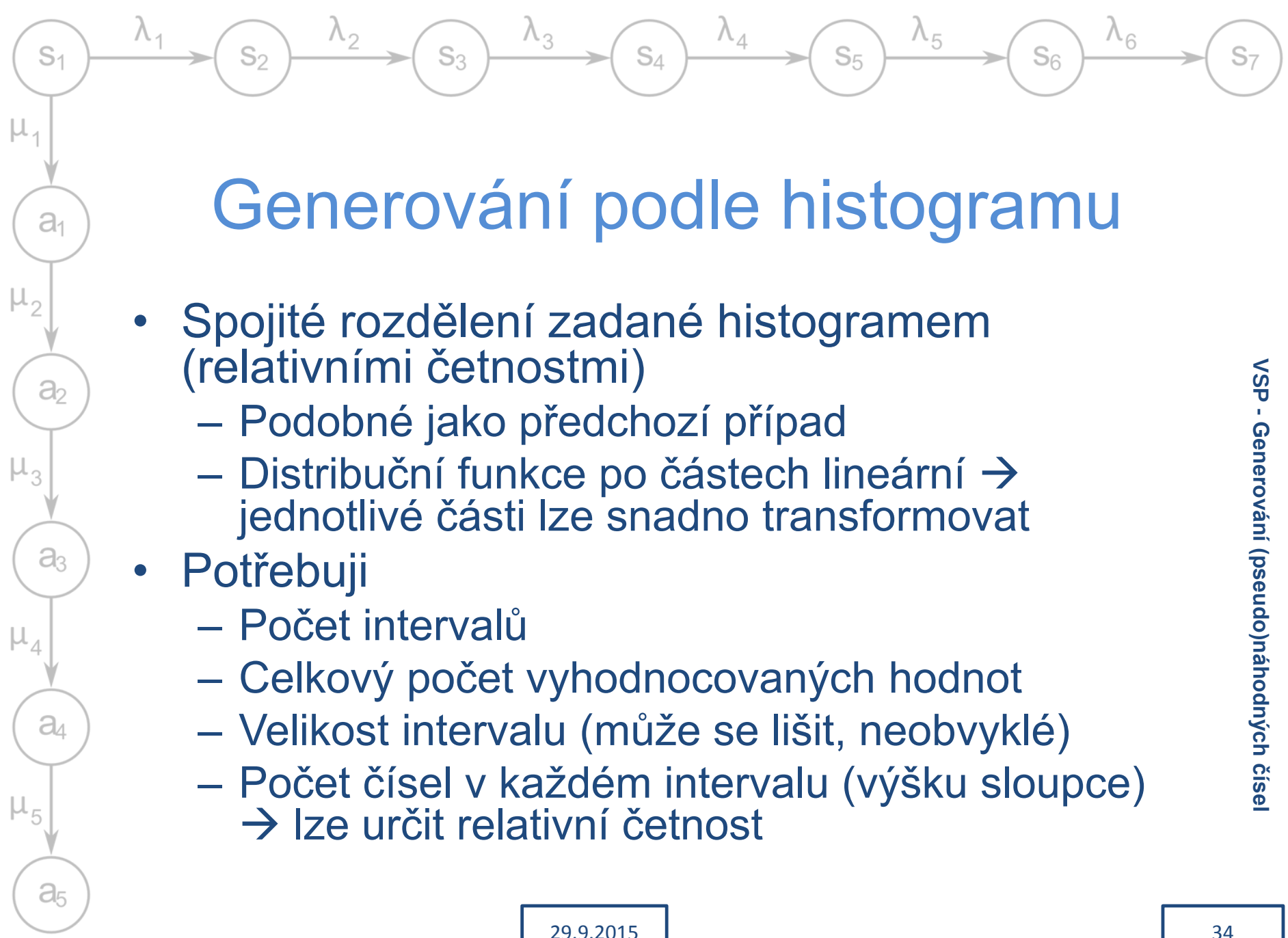
Obecné diskrétní rozdělení

- Několik hodnot se známou pravděpodobností (může být různá)
  - Např. 1 – 30%, 2 – 50%, 3 – 20% (součet musí být 100%)

→„Schodová“ distribuční funkce

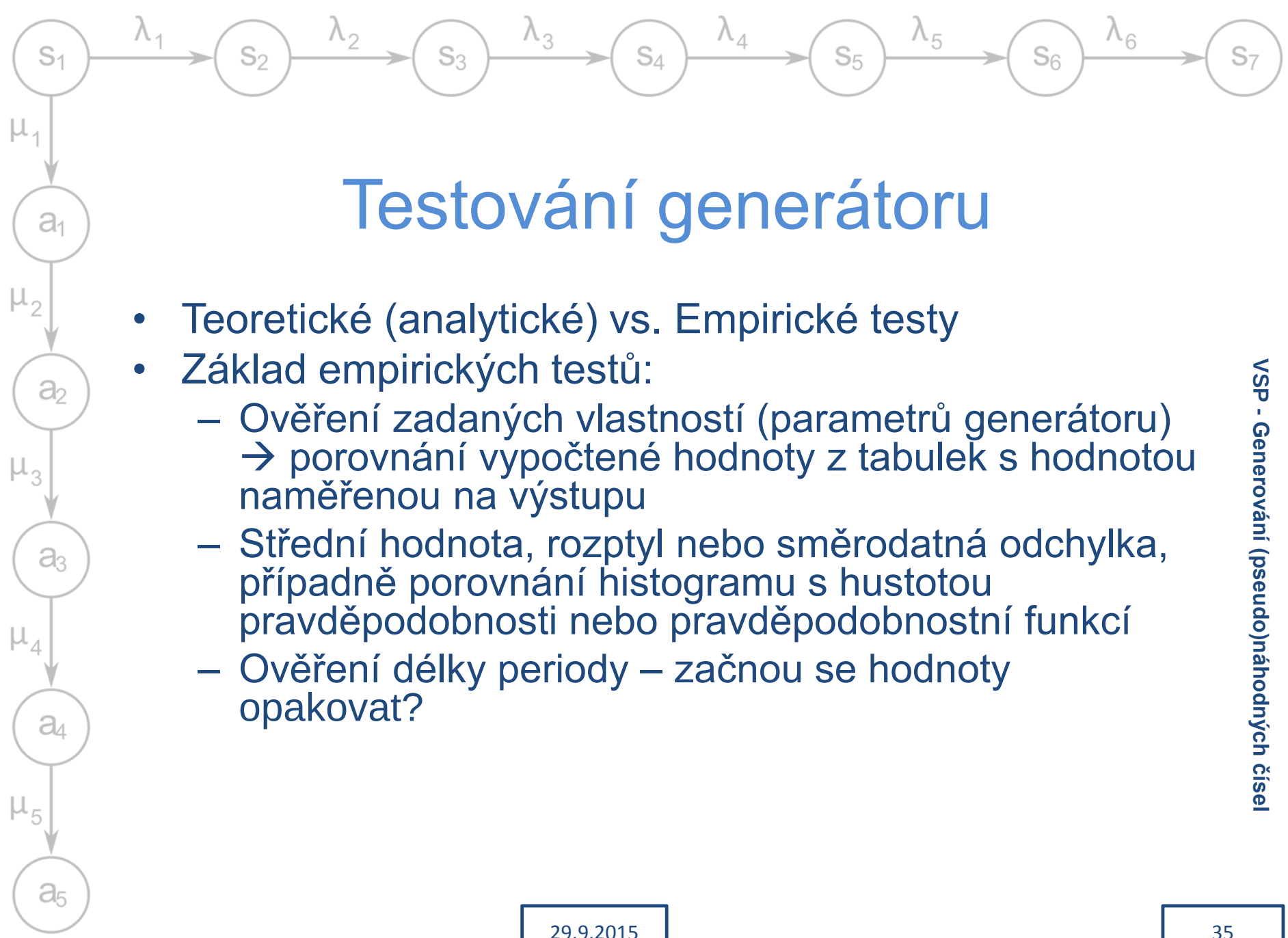
- Ekvivalent transformační metody
  - Generuji 1 náhodné číslo s normovaným uniformním rozdělením
  - Podle tabulky určím výslednou hodnotu

od	do	hodnota
0	0,3	1
0,3	0,8	2
0,8	1	3



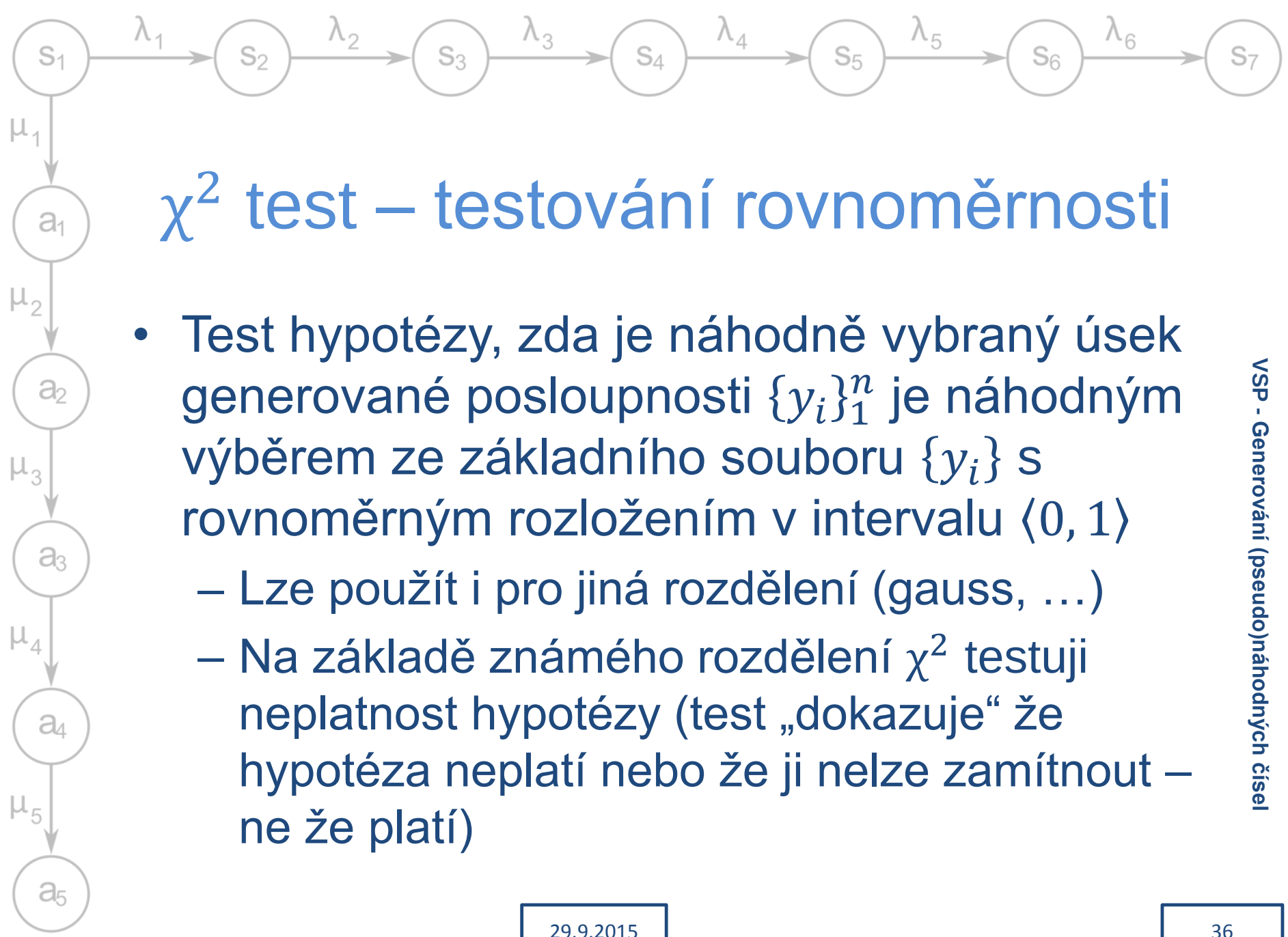
Generování podle histogramu

- Spojité rozdělení zadané histogramem (relativními četnostmi)
  - Podobné jako předchozí případ
  - Distribuční funkce po částech lineární → jednotlivé části lze snadno transformovat
- Potřebuji
  - Počet intervalů
  - Celkový počet vyhodnocovaných hodnot
  - Velikost intervalu (může se lišit, neobvyklé)
  - Počet čísel v každém intervalu (výšku sloupce) → lze určit relativní četnost



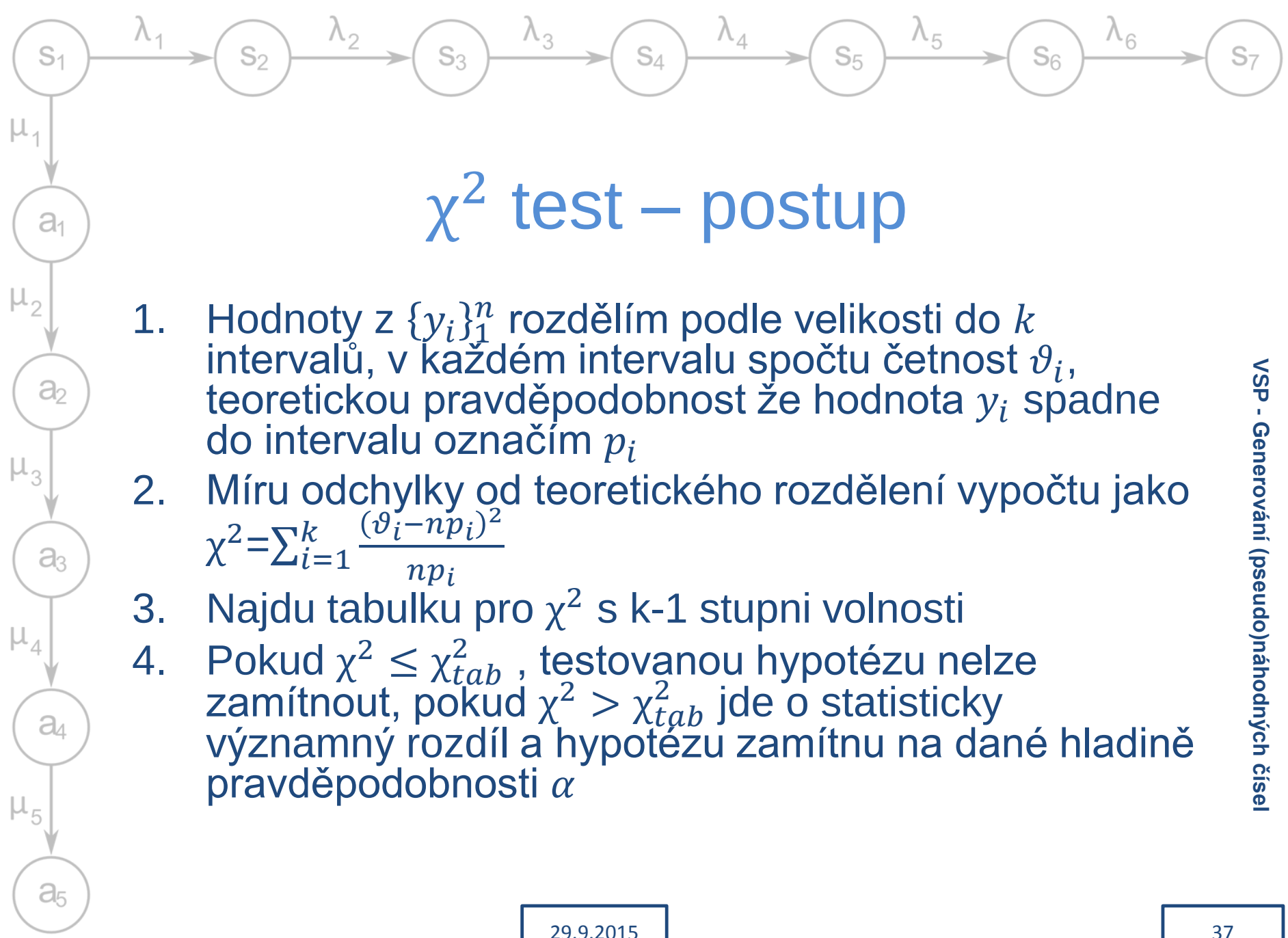
Testování generátoru

- Teoretické (analytické) vs. Empirické testy
- Základ empirických testů:
  - Ověření zadaných vlastností (parametrů generátoru) → porovnání vypočtené hodnoty z tabulek s hodnotou naměřenou na výstupu
  - Střední hodnota, rozptyl nebo směrodatná odchylka, případně porovnání histogramu s hustotou pravděpodobnosti nebo pravděpodobnostní funkcí
  - Ověření délky periody – začnou se hodnoty opakovat?



χ^2 test – testování rovnoměrnosti

- Test hypotézy, zda je náhodně vybraný úsek generované posloupnosti $\{y_i\}_1^n$ je náhodným výběrem ze základního souboru $\{y_i\}$ s rovnoměrným rozložením v intervalu $\langle 0, 1 \rangle$
  - Lze použít i pro jiná rozdělení (gauss, …)
  - Na základě známého rozdělení χ^2 testuji neplatnost hypotézy (test „dokazuje“ že hypotéza neplatí nebo že ji nelze zamítnout – ne že platí)

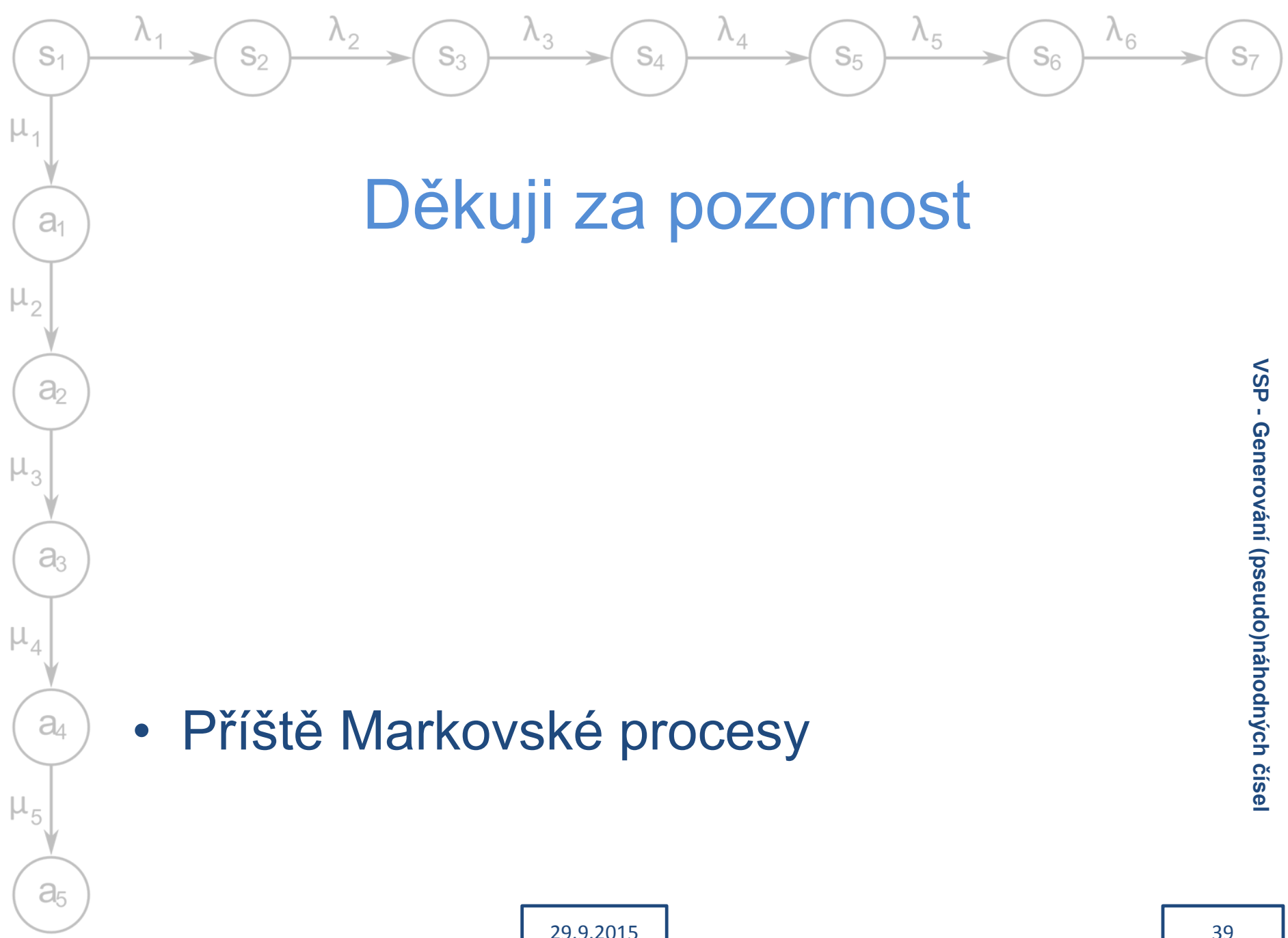


χ^2 test – postup

1. Hodnoty z $\{y_i\}_1^n$ rozdělím podle velikosti do k intervalů, v každém intervalu spočtu četnost ϑ_i, teoretickou pravděpodobnost že hodnota y_i spadne do intervalu označím p_i
2. Míru odchylky od teoretického rozdělení vypočtu jako $\chi^2=\sum_{i=1}^{k}\frac{(\vartheta_i-np_i)^2}{np_i}$
3. Najdu tabulku pro χ^2 s k-1 stupni volnosti
4. Pokud $\chi^2 \leq \chi^2_{tab}$, testovanou hypotézu nelze zamítnout, pokud $\chi^2 > \chi^2_{tab}$ jde o statisticky významný rozdíl a hypotézu zamítnu na dané hladině pravděpodobnosti α

Užitečné zdroje

- **Uncommons maths** - http://maths.uncommons.org/
  - Přesná matematika (`Rational`), kombinatorika, statistika, náhodná čísla
- **Random.org** - https://www.random.org/
  - Skutečná náhodná čísla online
- **NIST Computer security division** - http://csrc.nist.gov/groups/ST/toolkit/rng/index.html
  - Návody jak testovat kvalitu pseudonáhodných generátorů (standard)



Děkuji za pozornost

- Příště Markovské procesy