

Galtonova deska a náhodná procházka

Jiří Jičínský, Michal Jelínek, Tomáš Sláma

Náhodná procházka

- Dráha molekuly, cena akcií
- **Diskrétní** procházka (stejné kroky)
- Pravděpodobnost jako hod mincí – **50/50**
- Spojitost s **Pascalovým trojúhelníkem**

Krok/pozice	-4	-3	-2	-1	0	1	2	3	4
0 krok					1				
1 krok				1/2		1/2			
2 krok			1/4		1/2		1/4		
3 krok		1/8		3/8		3/8		1/8	
4 krok	1/16		1/4		3/8		1/4		1/16

Binomické rozdělení pro náhodnou procházku

Celkový počet kroků

Zde je zahrnuta pravděpodobnost z mince (50/50)

$$P(d) = \binom{N}{\frac{d+N}{2}} \frac{1}{2^N} = \frac{N!}{\left(\frac{N-d}{2}\right)! \left(\frac{N+d}{2}\right)!} \frac{1}{2^N}$$

Vzdálenost od začátku

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} - \text{kombinační číslo}$$

- Příklad: $(N = 12; d = 12) \dots P(12) = \binom{12}{12} \frac{1}{2^{12}} = 1 * \frac{1}{4096}$

Galtonova deska

- **Praktický** příklad náhodné procházky
- Dvě plexiskla – mezi jsou hroty simulující náhodný děj
- Data z testů by se měla **přibližovat** k binomickému rozdělení



χ^2 test dobré shody

- Test zjišťující **podobnost** našich dat s binomickým rozdělením
- Testuje nulovou hypotézu H_0 oproti alternativní hypotéze H_a

$$\chi^2 = \sum_{i=0}^{\text{pocet testovanych}} \frac{(\text{ocekavane}_i - \text{namerene}_i)^2}{\text{ocekavane}_i}$$

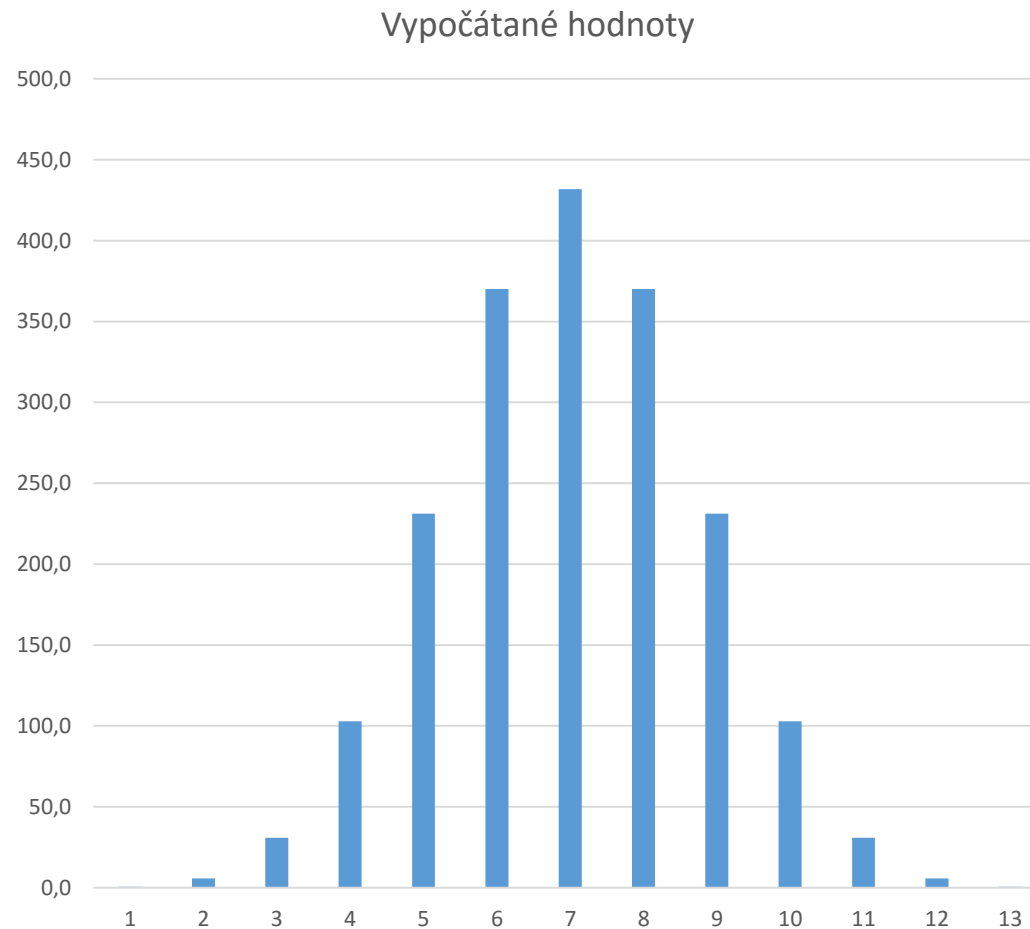
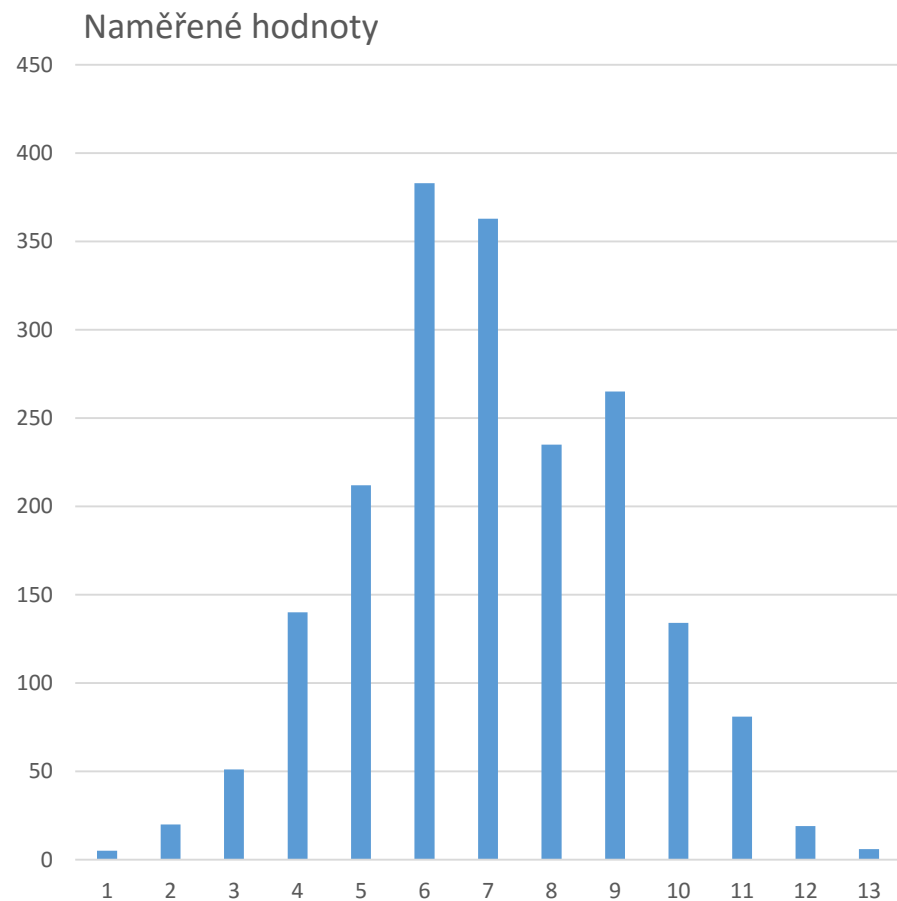
- Vyžaduje **velký** počet dat a min. teoretickou četnost 5 (v 80% případů)

Nasbíraná data

Pozice	-12	-10	-8	-6	-4	-2	0	2	4	6	8	10	12	Součet
Total	5	20	51	140	212	383	363	235	265	134	81	19	6	1914
Vypočítaná	0,5	5,6	30,8	102,8	231,3	370,1	431,8	370,1	231,3	102,8	30,8	5,6	0,5	1914
Chí suma	44,0	36,9	13,2	13,5	1,6	0,5	11,0	49,3	4,9	9,5	81,6	32,0	65,5	363,3

- Chí suma se porovná s **kritickou hodnotou** ze statistických tabulek (podle **stupňů volnosti** a **hladiny významnosti**, která je většinou 0,05)
- H_0 ... Naše Galtonova deska se řídí binomickým rozdělením s pravděpodobností $\frac{1}{2}$.
- H_a ... Naše Galtonova deska se tímto rozdělením neřídí.

Grafy z nasbíraných dat

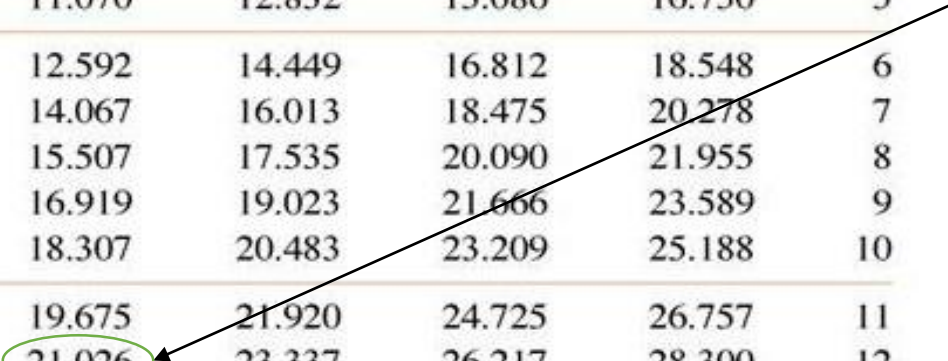


Tabulka kritických hodnot

- Pokud je naše chí suma **větší** než kritická hodnota, hypotézu H_0 zamítáme

df \ p	0.995	0.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005	df
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	12
13	3.565	5.009	7.042	12.340	19.812	22.362	24.736	27.688	29.819	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	15

Naše
kritická
hodnota



Výsledek

- Naše chí suma je **363,3**, kritická hodnota je **21,06**
- Hypotézu H_0 o binomickém rozdělení s pravděpodobností $\frac{1}{2}$ **zamítáme**
- Možně řešení H_1 ... Naše Galtonova deska se řídí binomickým rozdělením s jinou pravděpodobností než $\frac{1}{2}$
 - Pomocí dat jsme odhadli pravděpodobnost realizace kroku doprava s výsledkem 0,503
 - Odchylka oproti 0,5 je příliš malá, na velké chí-sumě by to nic nezměnilo

P	0	20	102	420	848	1915	2178	1645	2120	1206	810	209	72	0,503
----------	---	----	-----	-----	-----	------	------	------	------	------	-----	-----	----	--------------

Závěr

- **Nelze** říci, že se naše deska řídí binomickým rozdělením
- Možné důvody
 - Nepřesnost kolíků
 - Vysoká startovní rychlost
 - Nerovnost podstavy
 - Malé množství pokusů