

Fyzikální modely a jejich testování

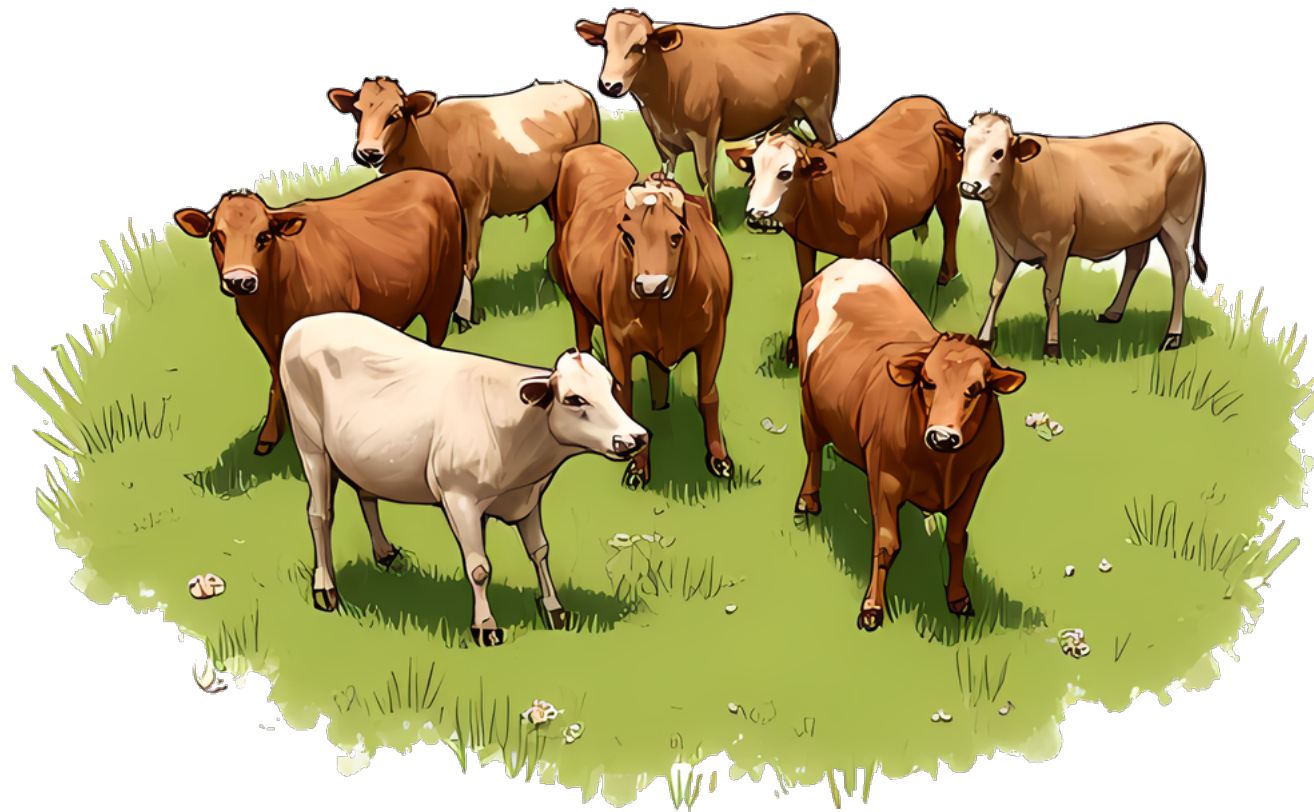


Karel Lemr (k.lemr@upol.cz)
<http://jointlab.upol.cz/lemr>, @KvantovyKarel

realita:

příliš složitá na popsání

mnoho vlastností není
předmětem zkoumání



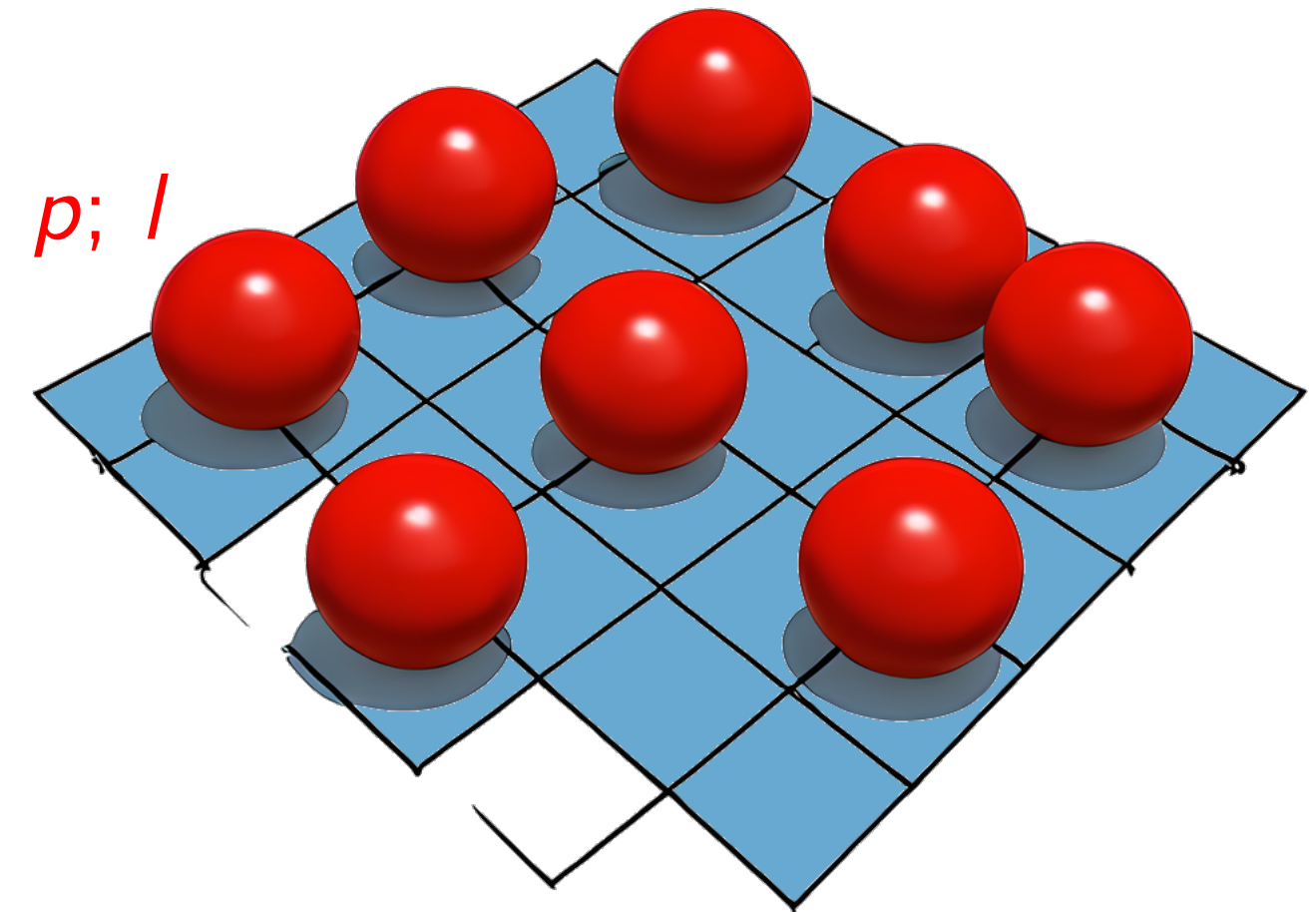
model:

zjednodušený popis

umožní popsat potřebné aspekty

vznikají nepřesnosti

$m; r; p; l$



fitování

experimentální data

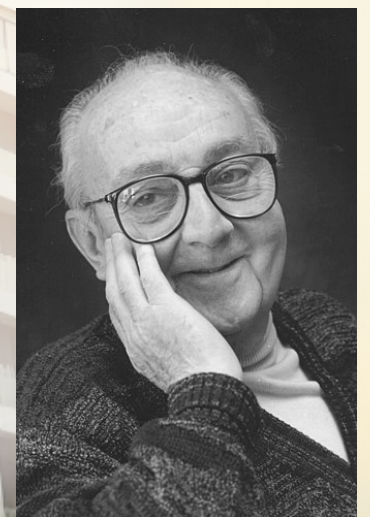


parametry modelu



**"All models are wrong,
but some are useful."**

-- George E. P. Box

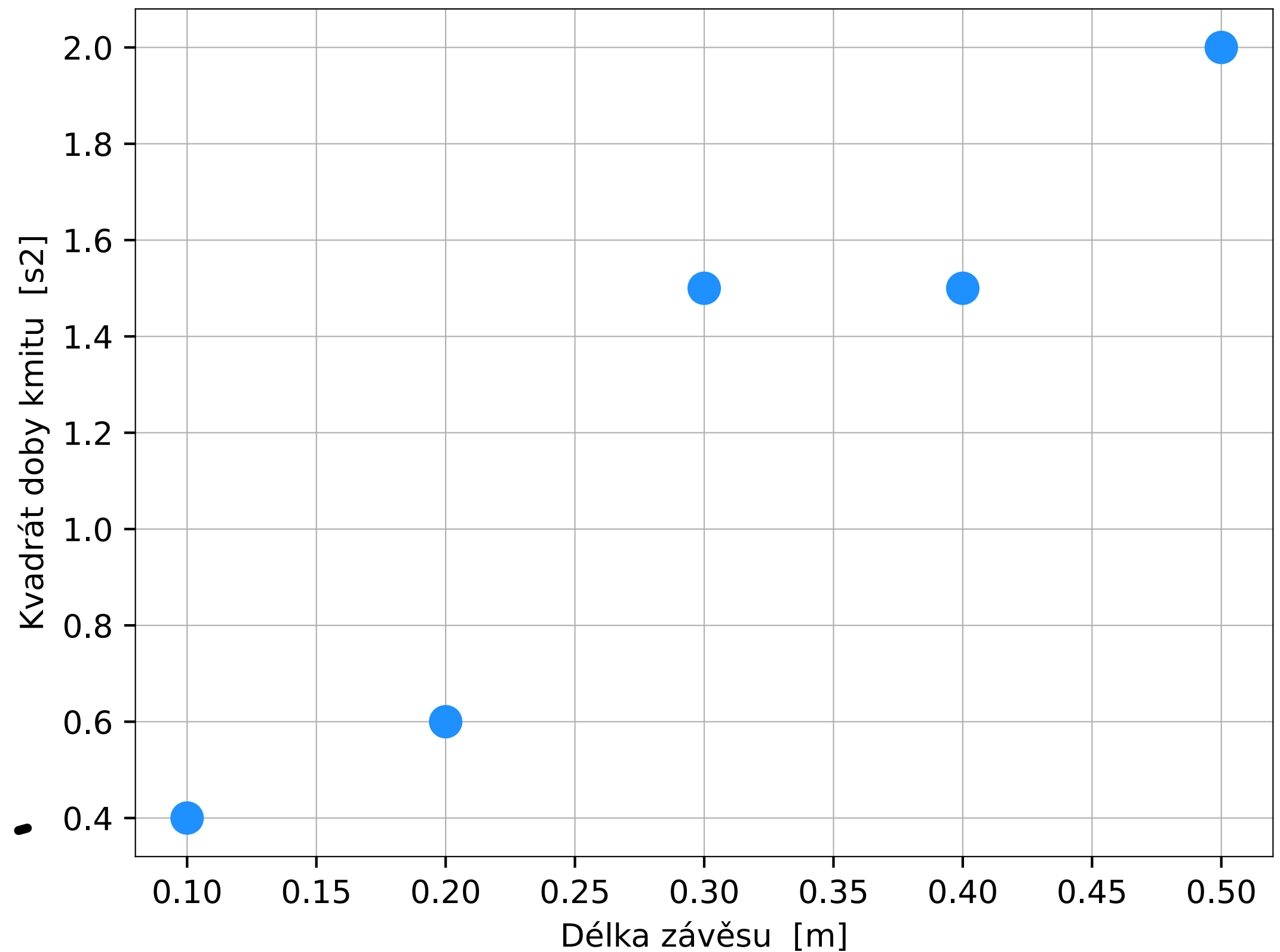


$$T^2 = \frac{4\pi^2}{g} d$$

doba kyvu T



délka d



experiment: měříme T pro různá d

model: pevný bod na nehmotném závěsu, infinitezimálně malá amplituda kyvu



Lineární regrese:

fitujeme funkci přímky $\hat{y} = ax + b$

hledáme parametry a, b

stříška označuje
hodnotu předpovězenou
modelem

Otázky:

1. jak najít a, b
2. jak dobře fit (model) pasuje
3. není fit (model) jen náhoda

ad 1. minimalizujeme:

$$\text{RES} = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$



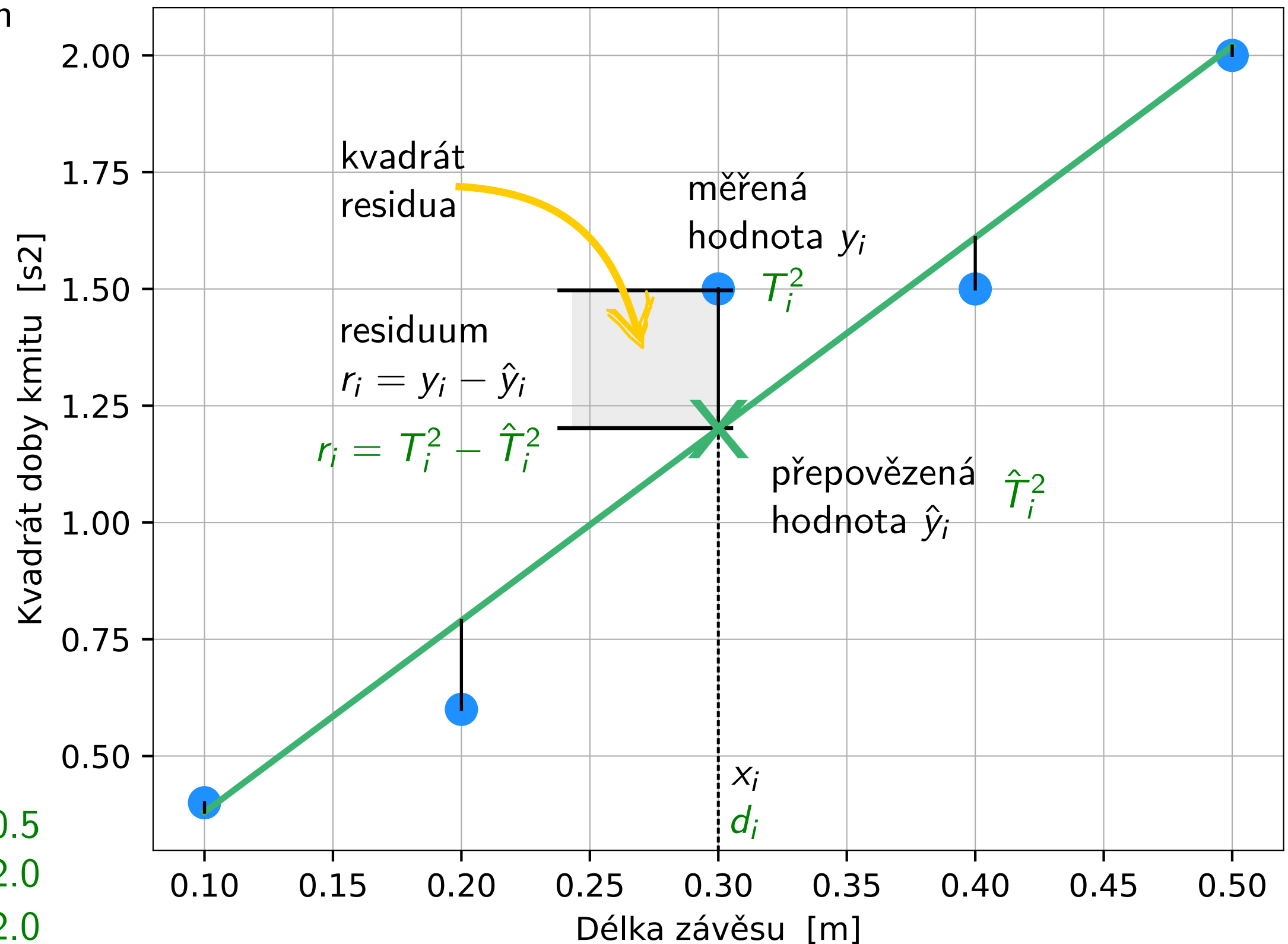
Existuje
analytické
řešení.

d [m]	0.1	0.2	0.3	0.4	0.5
T^2 [s ²]	0.4	0.6	1.5	1.5	2.0
$\hat{T}^2$ [s ²]	0.4	0.8	1.2	1.6	2.0

konkrétní příklad:

$$\hat{T}^2 = ad + 0$$

hledáme parametr a , následně $g = \frac{4\pi^2}{a}$



Lineární regrese:

fitujeme funkci přímky $\hat{y} = ax + b$
hledáme parametry a, b

pokud bychom neuvažovali
závislost na x, y má určitý rozptyl
(neurčitost) $SS = \sum_{i=1}^n (y_i - \bar{y})^2$

využitím modelu
(očekávané závislosti na x), klesne
neurčitost $RES = \sum_{i=1}^n (y_i - \hat{y}_i)^2$

ad 2. vypočteme:

$$R^2 = \frac{SS - RES}{SS}$$

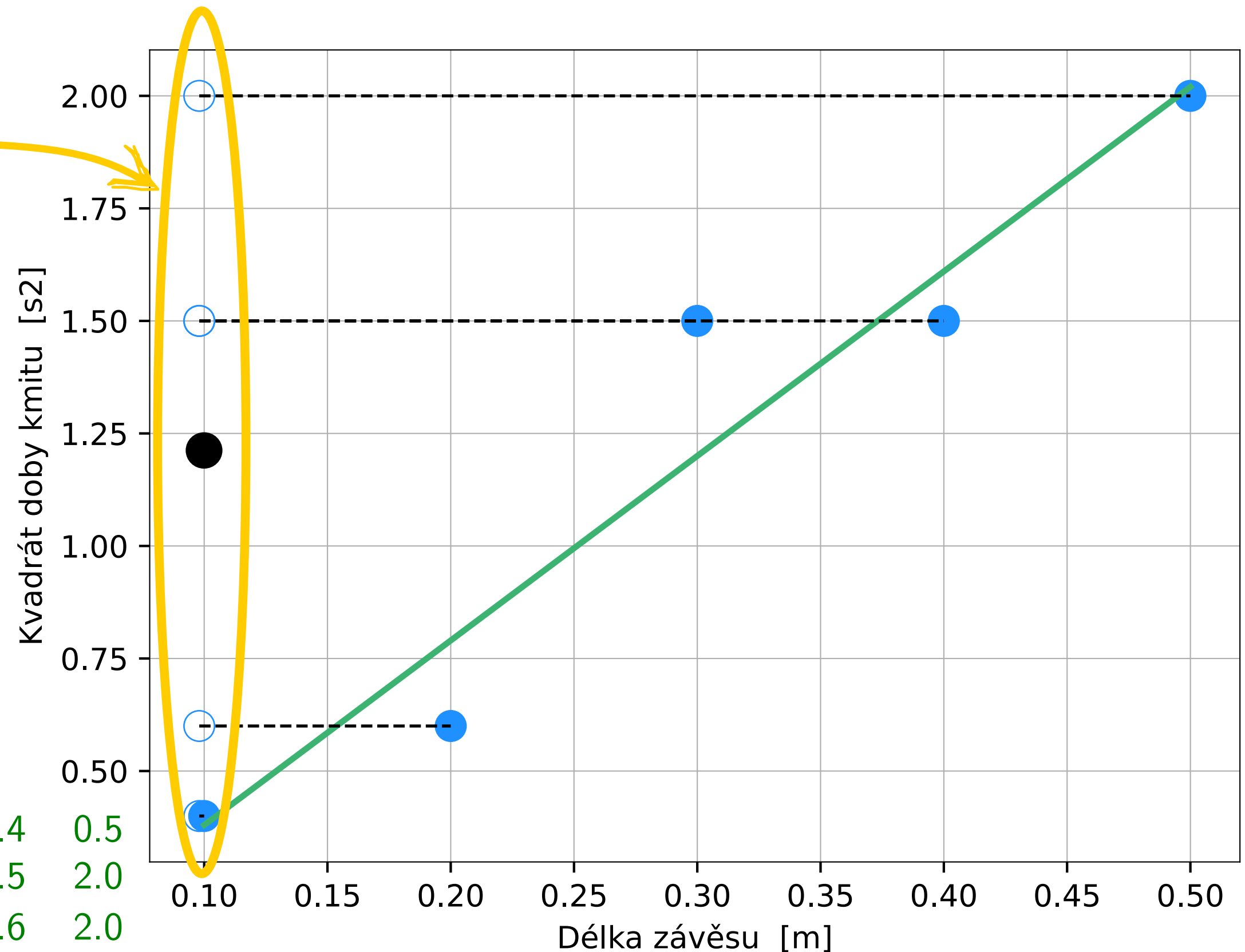
koeficient determinace

d [m]	0.1	0.2	0.3	0.4	0.5
T^2 [s ²]	0.4	0.6	1.5	1.5	2.0
$\hat{T}^2$ [s ²]	0.4	0.8	1.2	1.6	2.0

konkrétní příklad:

$$\hat{T}^2 = ad + 0$$

hledáme parametr a , následně $g = \frac{4\pi^2}{a}$



Lineární regrese:

fitujeme funkci přímky $\hat{y} = ax + b$
hledáme parametry a, b

pokud bychom neuvažovali
závislost na x , y má určitý rozptyl
(neurčitost) $SS = \sum_{i=1}^n (y_i - \bar{y})^2$

využitím modelu
(očekávané závislosti na x), klesne
neurčitost $RES = \sum_{i=1}^n (y_i - \hat{y}_i)^2$

ad 3. F-test:

$df_{\text{model}} = n - p$ p počet parametrů modelu

$df_{\text{bez}} = n - 1$

$df_j = df_{\text{model}}$

$df_c = df_{\text{bez}} - df_{\text{model}}$

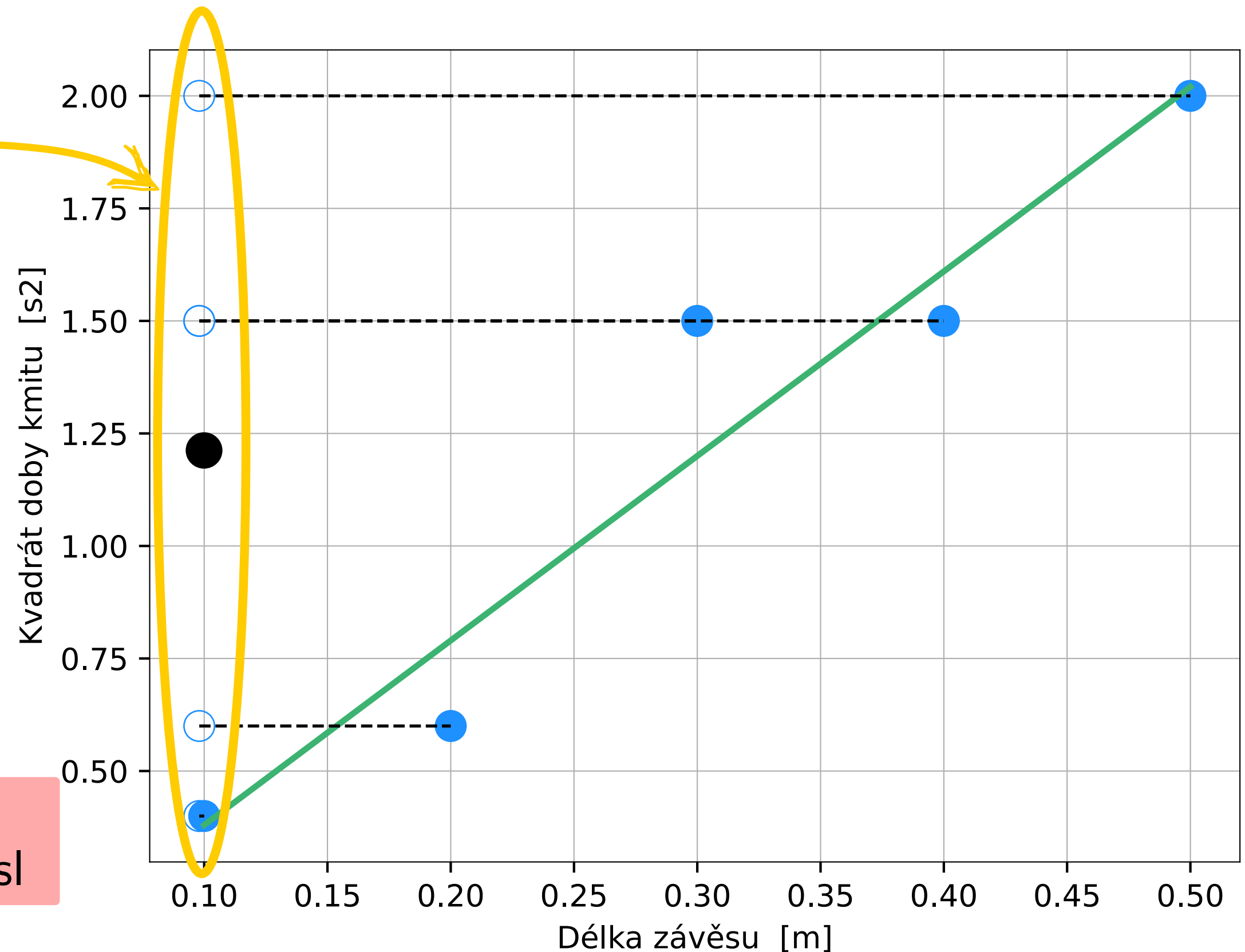
$$F = \frac{(SS - RES) / df_c}{RES / df_j}$$

$F > F_{\text{krit}}$
model má smysl

konkrétní příklad:

$$\hat{T}^2 = ad + 0$$

hledáme parametr a , následně $g = \frac{4\pi^2}{a}$



Tabulka kritických F hodnot pro hladinu 0,05

<https://www.oreilly.com/library/view/making-sense-of/9780470074718/appa-secoo4.html>

df _{jmenovatel} ↓	df _{čitatel} →																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88



ad 3. F-test:

$df_{\text{model}} = n - p$ p počet parametrů modelu

$df_{\text{bez}} = n - 1$

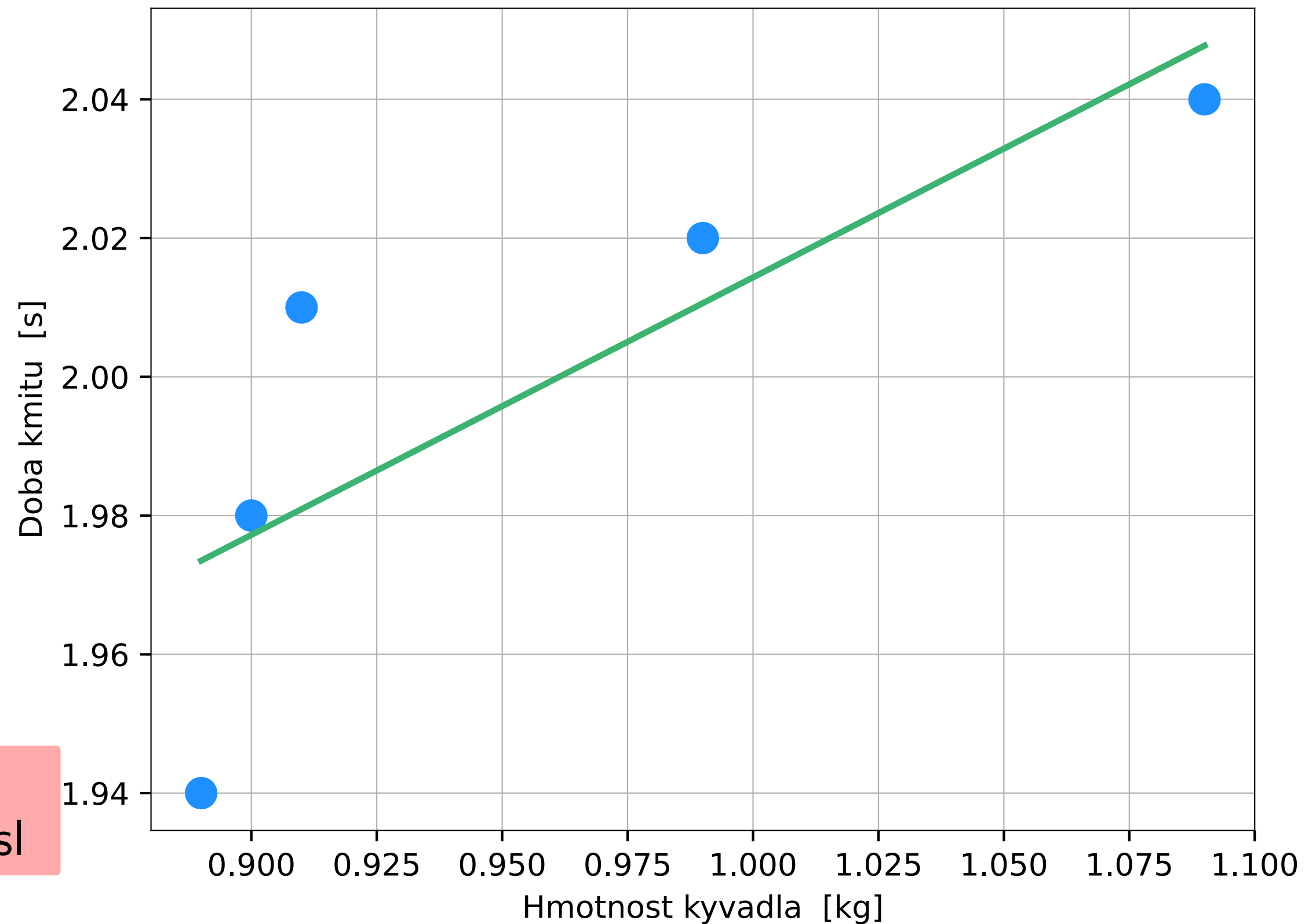
$df_j = df_{\text{model}}$

$df_c = df_{\text{bez}} - df_{\text{model}}$

$F = \frac{(SS - RES) / df_c}{RES / df_j}$

$F > F_{\text{krit}}$
model má smysl

$T = am + b$
 $SS = 0,0061$
 $RES = 0,0020$



Porovnání modelů:

Našla jsem
nový super
model!

Něco se
mi na tom
nezdá.

$$df_1 = n - p_1$$

$$df_2 = n - p_2$$

$$df_j = df_2$$

$$df_c = df_1 - df_2$$

$$F = \frac{(RES_1 - RES_2) / df_c}{RES_2 / df_j}$$

RES_1 ... suma čtverců residuí jednoduššího modelu
 df_1 .. stupně volnosti jednoduššího modelu ($n - p_1$)

RES_2 ... suma čtverců residuí složitějšího modelu
 df_2 .. stupně volnosti jednoduššího modelu ($n - p_2$)

