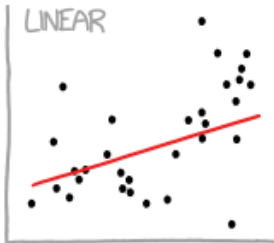
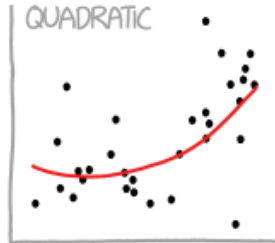


χ^2 test kvality fitu

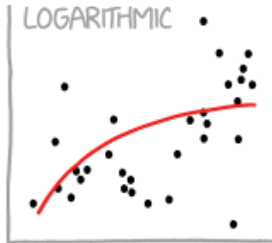
CURVE-FITTING METHODS AND THE MESSAGES THEY SEND



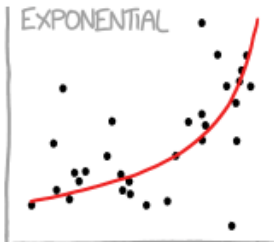
"HEY, I DID A REGRESSION."



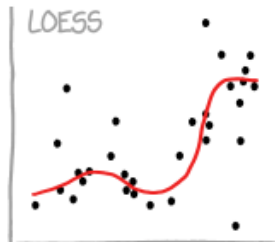
"I WANTED A CURVED LINE, SO I MADE ONE WITH MATH."



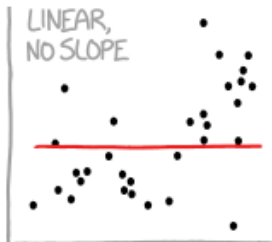
"LOOK, IT'S TAPERING OFF!"



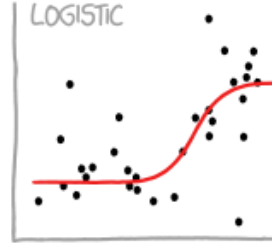
"LOOK, IT'S GROWING UNCONTROLLABLY!"



"I'M SOPHISTICATED, NOT LIKE THOSE BUMBLING POLYNOMIAL PEOPLE."



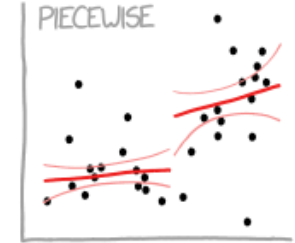
"I'M MAKING A SCATTER PLOT BUT I DON'T WANT TO."



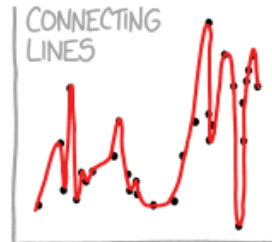
"I NEED TO CONNECT THESE TWO LINES, BUT MY FIRST IDEA DIDN'T HAVE ENOUGH MATH."



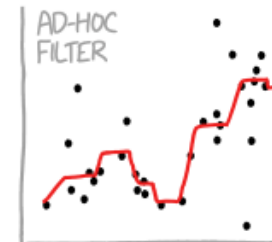
"LISTEN, SCIENCE IS HARD. BUT I'M A SERIOUS PERSON DOING MY BEST."



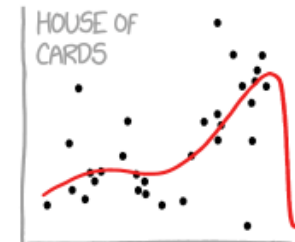
"I HAVE A THEORY, AND THIS IS THE ONLY DATA I COULD FIND."



"I CLICKED 'SMOOTH LINES' IN EXCEL."



"I HAD AN IDEA FOR HOW TO CLEAN UP THE DATA. WHAT DO YOU THINK?"



"AS YOU CAN SEE, THIS MODEL SMOOTHLY FITS THE- WAIT NO NO DON'T EXTEND IT AAAAAA!!!"

χ^2 test kvality fitu

- sada naměřených hodnot $\mathbf{x} = (x_1, x_2, \dots, x_n)$ (nezávislé proměnné)
 $\mathbf{y} = (y_1, y_2, \dots, y_n)$ (závislé proměnné) $y_i \in N(\mu_i, \sigma_i)$
- modelová funkce $\lambda(x|\boldsymbol{\theta})$ (modelujeme závislost $y(x)$)
 $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$ (parametry modelové závislosti)
- náhodná proměnná χ^2 $\chi^2(\boldsymbol{\theta}|\mathbf{y}, \boldsymbol{\sigma}, \mathbf{x}) = \sum_{i=1}^n \frac{(y_i - \lambda(x_i|\boldsymbol{\theta}))^2}{\sigma_i^2}$
- testovací statistika $\chi^2 \in f_{\chi^2}(\nu)$ χ^2 rozdělení o $\nu = n - m$ stupních volnosti

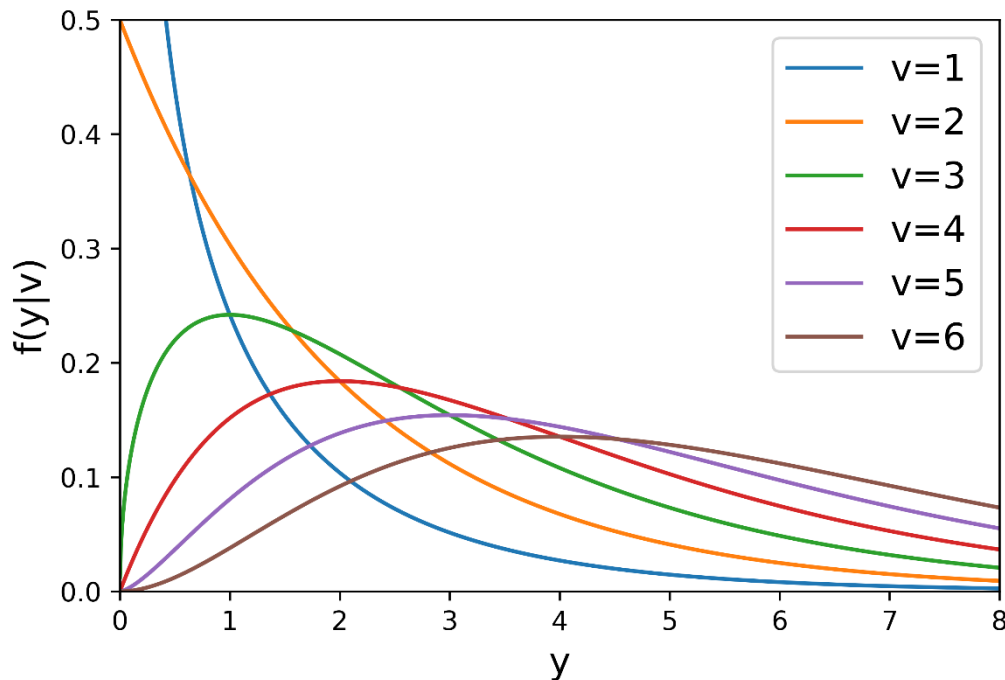
χ^2 rozdělení

- hustota pravděpodobnosti

$$f(y|\nu) = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} y^{\frac{\nu}{2}-1} e^{-\frac{y}{2}} \quad y \in [0, \infty) \quad \nu = 1, 2, \dots$$

počet stupňů volnosti ν

gama funkce $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$



- χ^2 rozdělení

$$z_i \in N(0,1)$$

$$y = \sum_{i=1}^{\nu} z_i^2 \rightarrow y \in f(\nu) \text{ resp. } \chi^2(\nu)$$

$$x_i \in N(\mu_i, \sigma_i)$$

$$y = \sum_{i=1}^{\nu} \frac{(x_i - \mu_i)^2}{\sigma_i^2} \rightarrow y \in f(\nu) \text{ resp. } \chi^2(\nu)$$

- momenty χ^2 rozdělení

$$E[y] = \nu$$

$$V[y] = 2\nu$$

χ^2 test kvality fitu

- χ^2 rozdělení $f(y|n - m) = \frac{1}{2^{\frac{n-m}{2}} \Gamma\left(\frac{n-m}{2}\right)} y^{\frac{n-m}{2}-1} e^{-\frac{y}{2}}$

$$P[y \geq \chi_0^2] = \int_{\chi_0^2}^{\infty} f(y|n - m) dy$$

- počet stupňů volnosti $\nu = n - m$

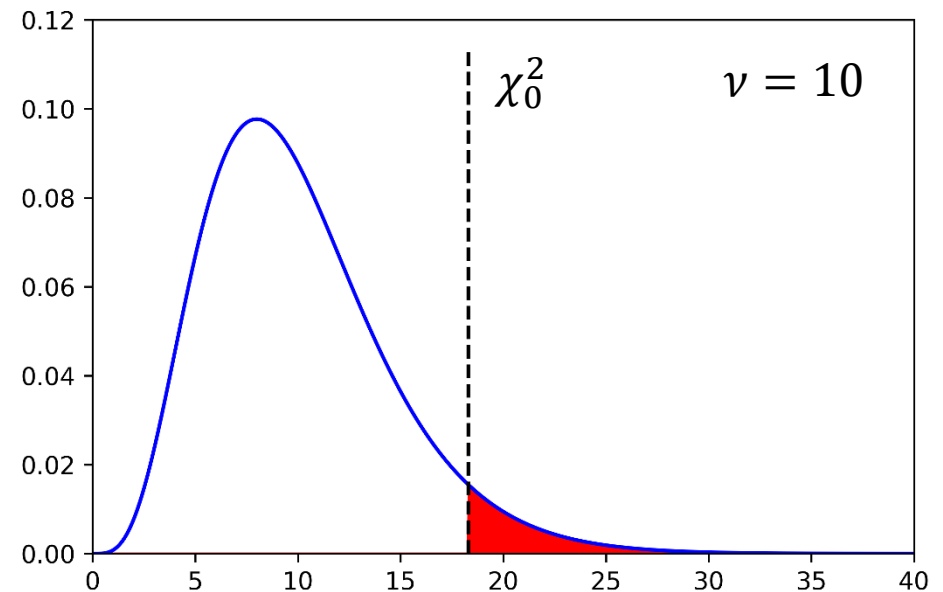
$$E[\chi^2] = n - m$$

$$V[\chi^2] = 2(n - m)$$

- χ^2 na počet stupňů volnosti $\chi^2/(n - m)$

$$E[\chi^2/(n - m)] = 1$$

$$V[\chi^2/(n - m)] = 2/(n - m)$$



- nulová hypotéza** H_0 naměřené hodnoty $y_i \in N(\mu_i, \sigma_i)$ jsou navzájem nezávislé a modelová funkce $\lambda(x, \theta)$ správně vystihuje závislost $y(x)$
- pokud je **pravděpodobnost** $P[y \geq \chi_0^2] < \alpha$, potom zamítneme nulovou hypotézu
- α **hladina signifikance** (typicky 0.05 nebo 0.01)

χ^2 test kvality fitu

- tabulka hodnot $P[y \geq \chi_0^2]$ pro počet stupňů volnosti $\nu = 1 - 10$

Počet stupňů volnosti ν	χ_0^2										
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88
10	3.94	4.87	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59
$P[y \geq \chi_0^2]$	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001

Pro počet stupňů volnosti $\nu > 10$ rozdělení $\chi^2(\nu)$ konverguje k normálnímu rozdělení $N(\nu, \sqrt{2\nu})$.

χ^2 test kvality fitu – polynom

- $m = 2$ ($\nu = 8$)

- $\chi^2 = 40.916$

- $\chi^2/\nu = 5.114$

- $0.001 < P$

- $m = 3$ ($\nu = 7$)

- $\chi^2 = 31.362$

- $\chi^2/\nu = 4.480$

- $0.001 < P$

- $m = 4$ ($\nu = 6$)

- $\chi^2 = 7.174$

- $\chi^2/\nu = 1.196$

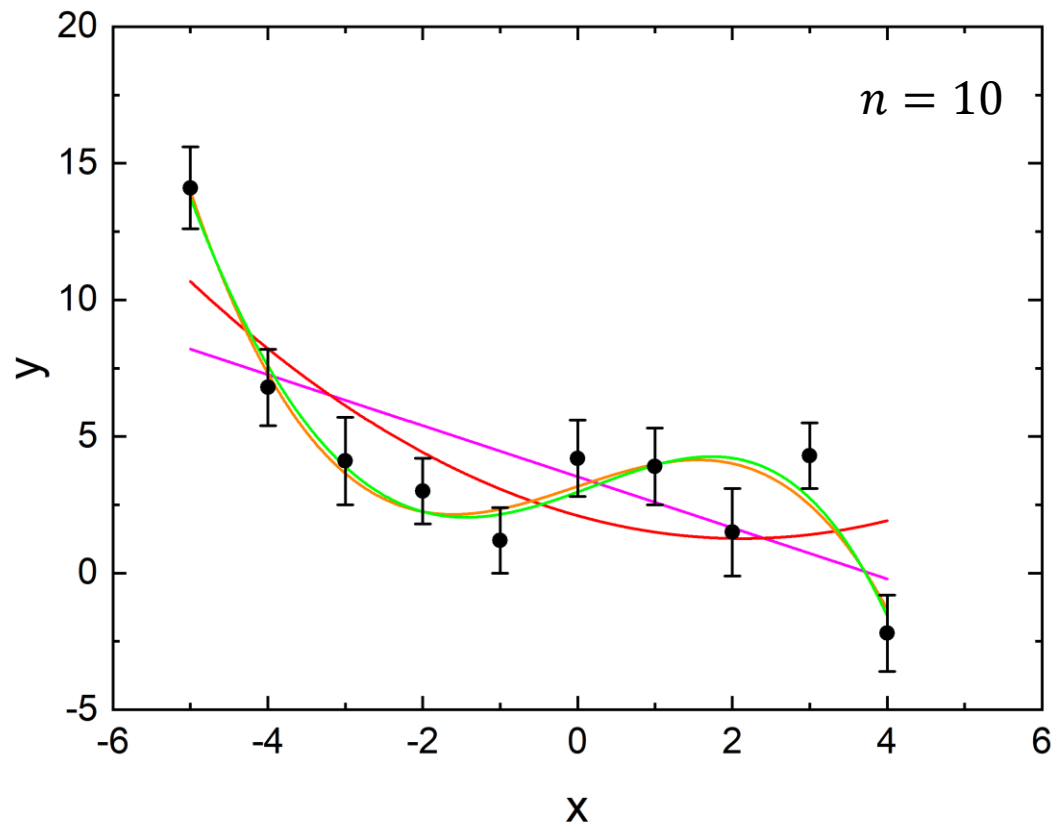
- $0.3 < P < 0.5$

- $m = 5$ ($\nu = 5$)

- $\chi^2 = 6.939$

- $\chi^2/\nu = 1.388$

- $0.2 < P < 0.3$



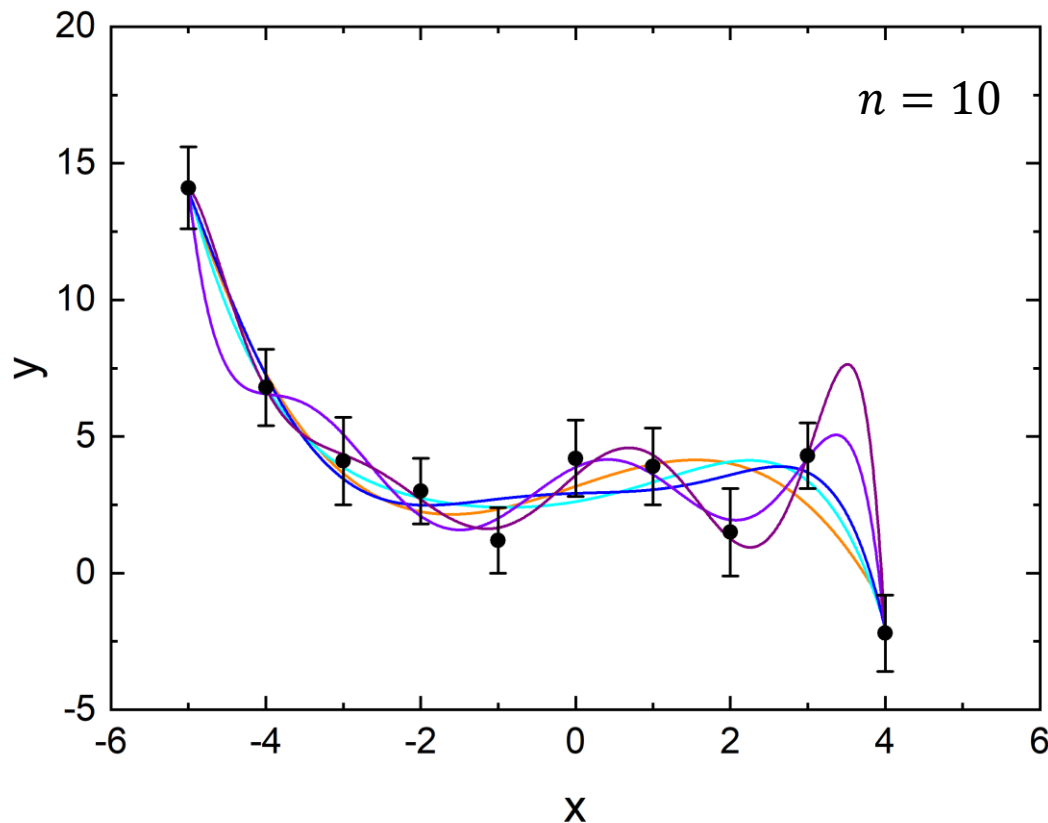
χ^2 test kvality fitu – polynom

- $m = 2$ ($\nu = 8$)
- $\chi^2 = 40.916$
- $\chi^2/\nu = 5.114$
- $0.001 < P$

- $m = 3$ ($\nu = 7$)
- $\chi^2 = 31.362$
- $\chi^2/\nu = 4.480$
- $0.001 < P$

- $m = 4$ ($\nu = 6$)
- $\chi^2 = 7.174$
- $\chi^2/\nu = 1.196$
- $0.3 < P < 0.5$

- $m = 5$ ($\nu = 5$)
- $\chi^2 = 6.939$
- $\chi^2/\nu = 1.388$
- $0.2 < P < 0.3$



- $m = 6$ ($\nu = 4$)
- $\chi^2 = 5.756$
- $\chi^2/\nu = 1.439$
- $0.2 < P < 0.3$

- $m = 7$ ($\nu = 3$)
- $\chi^2 = 5.230$
- $\chi^2/\nu = 1.743$
- $0.1 < P < 0.2$

- $m = 8$ ($\nu = 2$)
- $\chi^2 = 1.616$
- $\chi^2/\nu = 0.808$
- $0.3 < P < 0.5$

- $m = 9$ ($\nu = 1$)
- $\chi^2 = 0.545$
- $\chi^2/\nu = 0.545$
- $0.3 < P < 0.5$