Střední hodnota náh. veličiny

a) diskrétní

$$E(X) = \sum_i x_i \cdot p_i$$

b) spojitá

běžné označení

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx \quad (= \mu)$$

ROZPTYL (DISPERZE, VARIANCE) NÁHODNÉ VELIČINY

Def. $D(X) = E(X - E(X))^2 \; (= \sigma^2)$ ozn.

a) diskrétní náh. prom.

$$D(X) = \sum_i (x_i - E(X))^2 \cdot p_i$$

b) spojitá náh. prom.

$$D(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

Příklad: 1.) Hrací kostka – 1 hod

$$E(X) = 3{,}5$$

$$D(X) = \frac{1}{6}\left[(1-3{,}5)^2 + (2-3{,}5)^2 + \cdots \right.$$

$$\left. \cdots + (6-3{,}5)^2\right] = 2{,}9$$

2.) Alternativní rozdělení

$$E(X) = p$$

$$D(X) = (1-p)^2 \cdot p + (0-p)^2 (1-p)$$

$$= p(1-p)$$

3.) Binomické rozdělení pravděpodobnosti

$$E(X) = n.p$$

$$D(X) = \sum_{k=0}^{n} (k - np)^2 \binom{n}{k} p^k (1-p)^{n-k} =$$

$$= \dots = n.p.(1-p)$$

4.) Poisson

$$E(X) = \lambda$$

$$D(X) = \sum_{k=0}^{\infty} (k-\lambda)^2 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \dots$$

$$\dots = \lambda$$

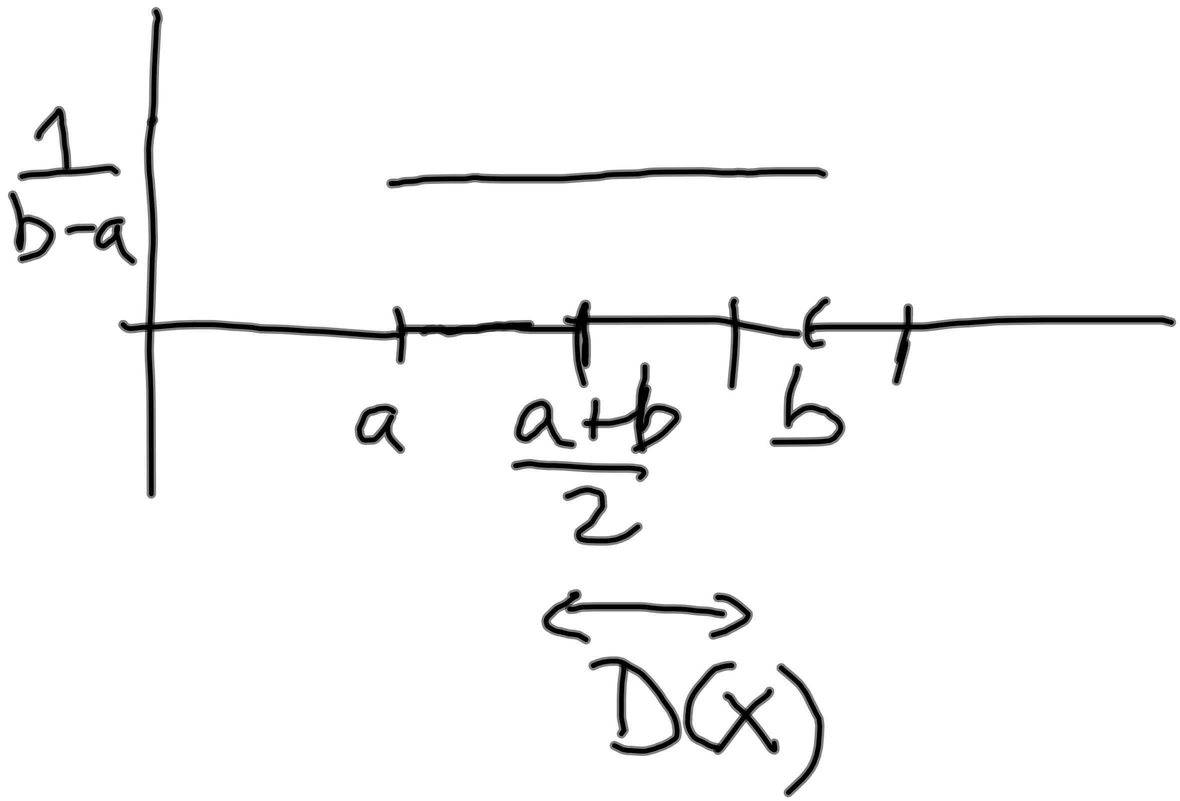
5.) Rovnoměrné rozdělení

$$E(X) = \frac{a+b}{2}$$

$$D(X) = \int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a}\, dx$$

$$x \in \langle a, b \rangle$$

$$D(X) = \int_{a}^{b} \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a}\, dx = \frac{(b-a)^2}{12}$$

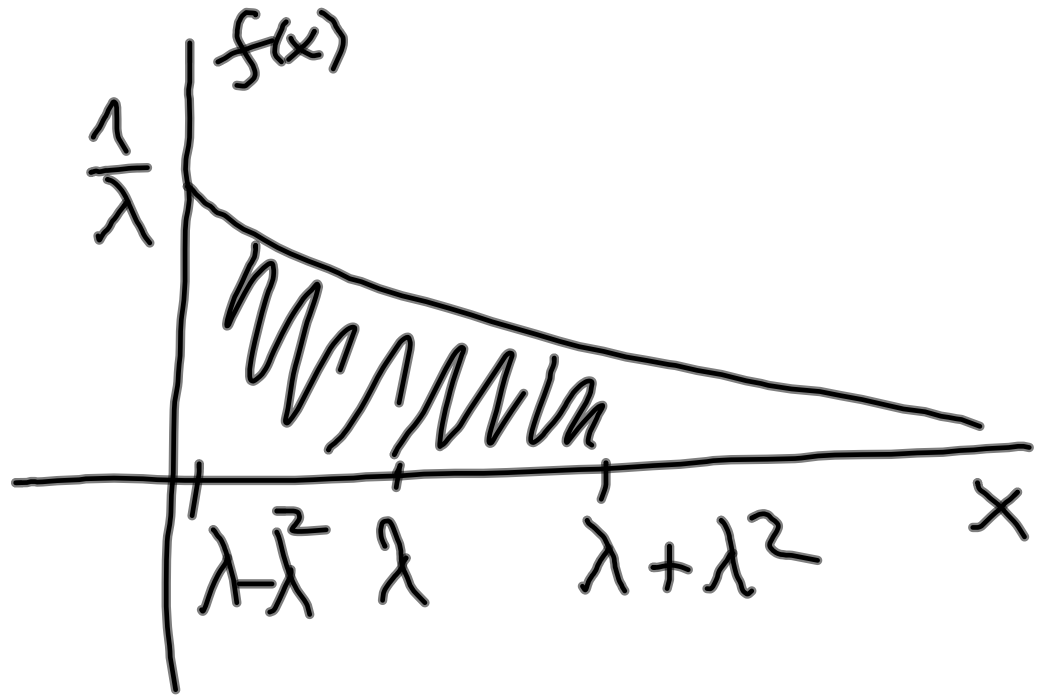

$\frac{1}{b-a}$
a
$\frac{a+b}{2}$
b
D(x)

6.) Exponenciální rozdělení

$$f(x) = \frac{1}{\lambda} e^{-\lambda x}, \quad \lambda > 0, x \geq 0$$

$$E(X) = \lambda$$

$$D(X) = \int_0^\infty (x-\lambda)^2 \cdot \frac{1}{\lambda} \cdot e^{-\lambda x} dx = \lambda^2$$


f(x)
1/λ
λ-λ²
λ
λ+λ²
x

7.) Normální (Gaussovo) rozdělení pravděpodobnosti

$$E(X) = \mu$$

$$D(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$

$$\ldots = \sigma^2$$

ČEBYŠEVOVA NEROVNOST

__Věta__. Nechť náh. prom. X má rozptyl $D(X)$. Pak pro libovolné $\varepsilon > 0$ je

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}$$

Důkaz. Pro spojitou náh. proměn.

$$D(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \geq$$

$$\geq \int_{-\infty}^{\mu-\varepsilon} \cdots + \int_{\mu+\varepsilon}^{\infty} \cdots \geq \varepsilon^2 \int_{-\infty}^{\mu-\varepsilon} f(x)\,dx +$$

$$+ \varepsilon^2 \int_{\mu+\varepsilon}^{\infty} f(x)\,dx = \varepsilon^2 P(|X - E(X)| \geq \varepsilon)$$

Teď konkrétně:

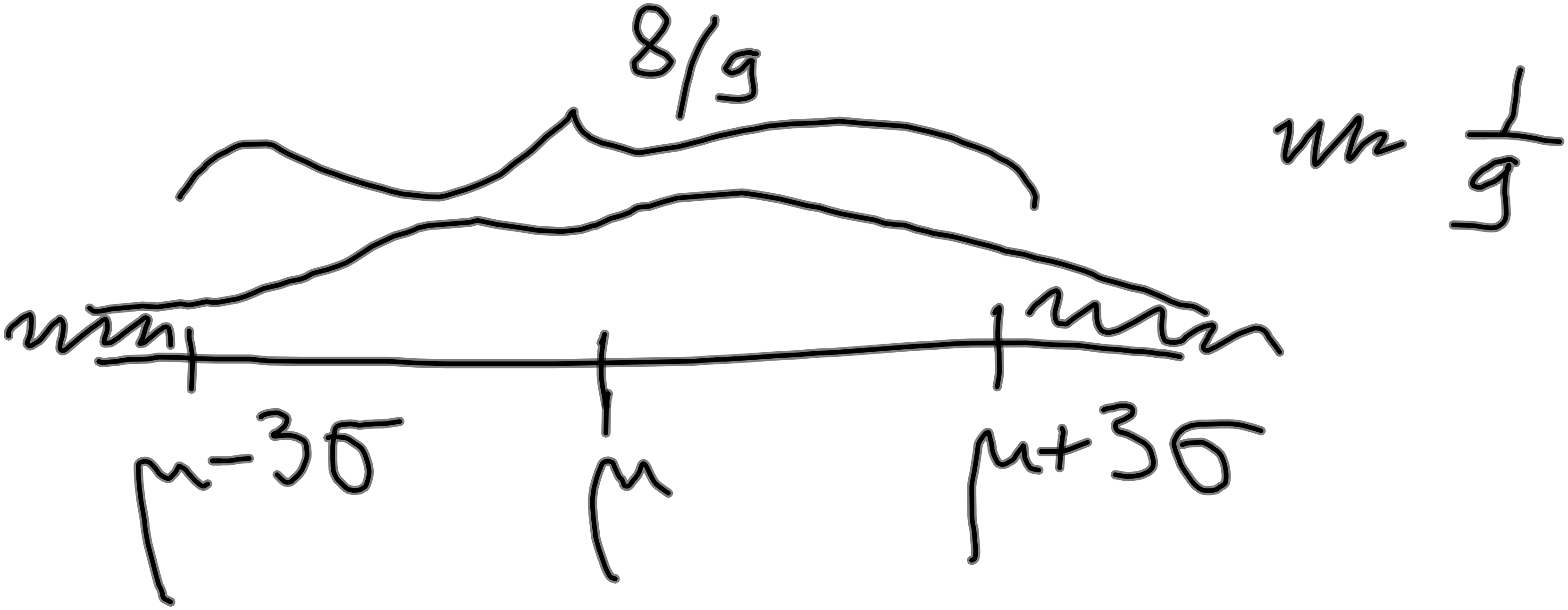
1.) $\varepsilon = \sigma$ (střední kvadratická odchylka)

$$P(|X-\mu| \geq \sigma) \leq \frac{\sigma^2}{\sigma^2} = 1$$

2.) $\varepsilon = 3\sigma$

$$P(|X-\mu| \geq 3\sigma) \leq \frac{\sigma^2}{(3\sigma)^2} = \frac{1}{9}$$

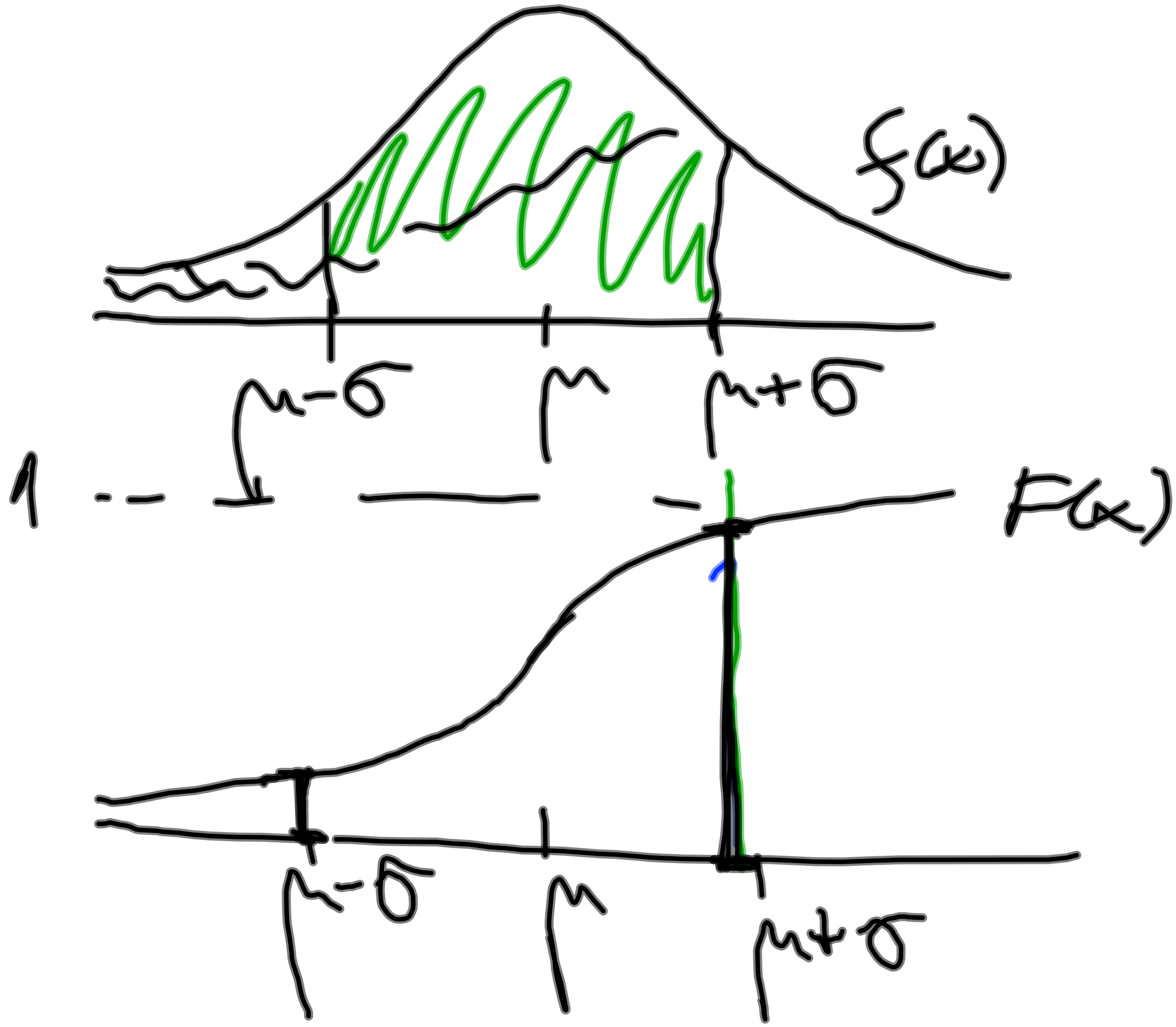
$$P(|X-\mu| < 3\sigma) \geq \frac{8}{9}$$


8/9
1/9
μ-3σ
μ
μ+3σ

Kdybychom věděli, že X má normální rozdělení $N(\mu, \sigma^2)$

$$P(|X-\mu|<3\sigma) =$$
$$= P(\mu-3\sigma < X < \mu+3\sigma) =$$
$$= F(\mu+3\sigma) - F(\mu-3\sigma) = \cdots$$

transformace na $N(0,1)$



f(x)
μ-σ
μ
μ+σ
1
F(x)
μ-σ
μ
μ+σ

Označíme-li distribuční funkci normalizovaného Gaussova rozdělení $\Phi(x)$, pak je

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

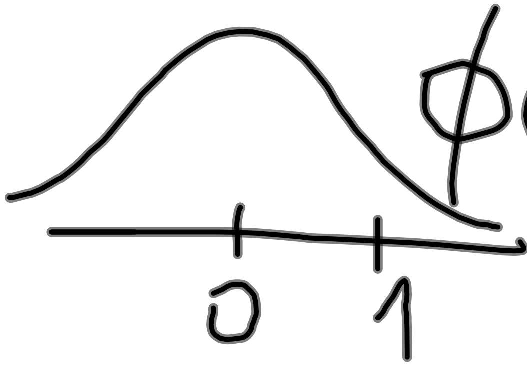
Proč?: $\displaystyle\int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} ds = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$

$$\frac{s-\mu}{\sigma} = t$$

Pozorování ... $F(\mu+3\sigma)-F(\mu-3\sigma)=$

$$=\phi\left(\frac{\mu+3\sigma-\mu}{\sigma}\right)-\phi\left(\frac{\mu-3\sigma-\mu}{\sigma}\right)=$$

$$=\phi(3)-\phi(-3)=\phi(3)-1+\phi(3)$$



$$\phi(-x)=1-\phi(x)$$

$$=2\phi(3)-1$$

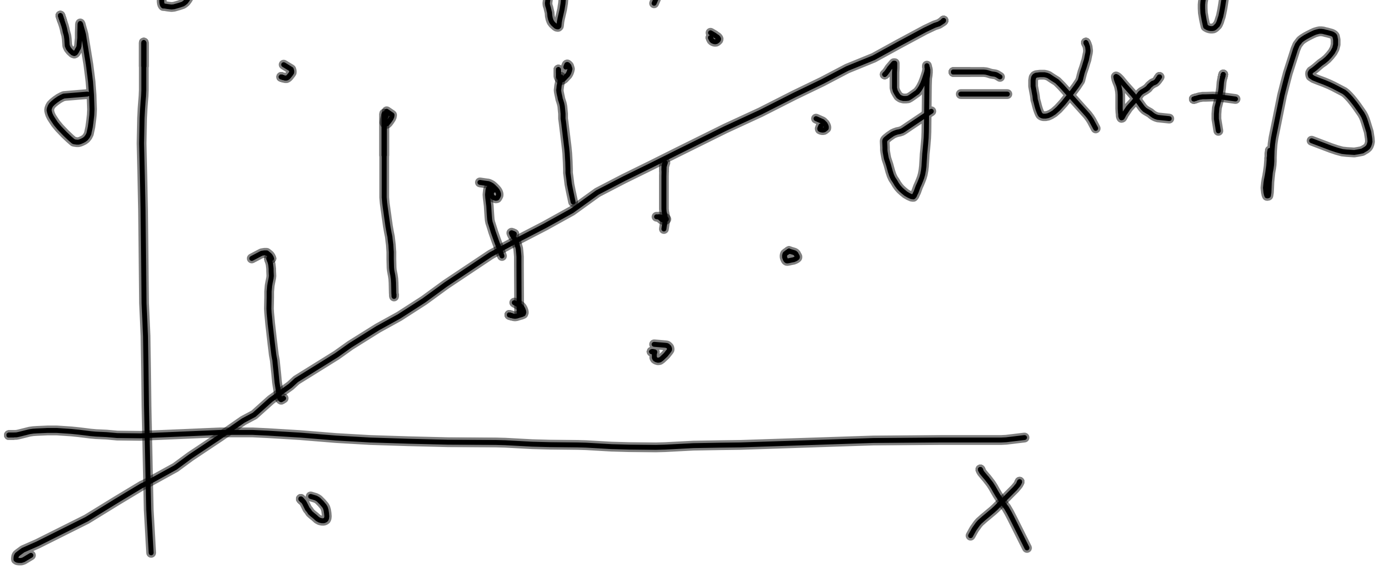
$$=2\cdot 0{,}99865-1$$

$$=0{,}9973$$

REGRESNÍ PŘÍMKY A KORELAČNÍ KOEFICIENT

n bodů

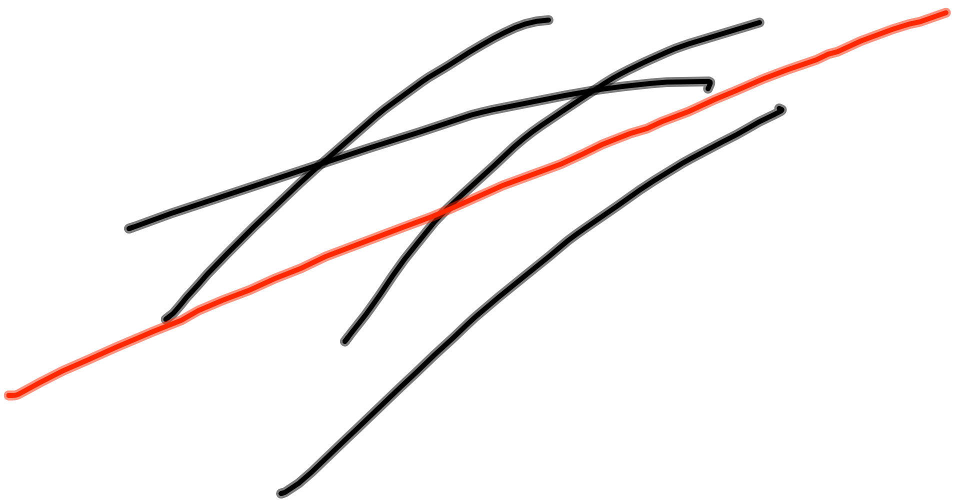
$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$



Jak se určují α, β?

$$\sum_{i=1}^{n} (y_i - \alpha x_i - \beta)^2 \leq \sum_{i=1}^{n} (y_i - a x_i - b)^2$$

$a, b \in R$



Je třeba najít

$$\min\left(f(a,b)=\sum_{i=1}^{n}(y_i-ax_i-b)^2\right)$$

Platí: $\exists\, i,j : x_i \neq x_j$

Rovnice regresní přímky je

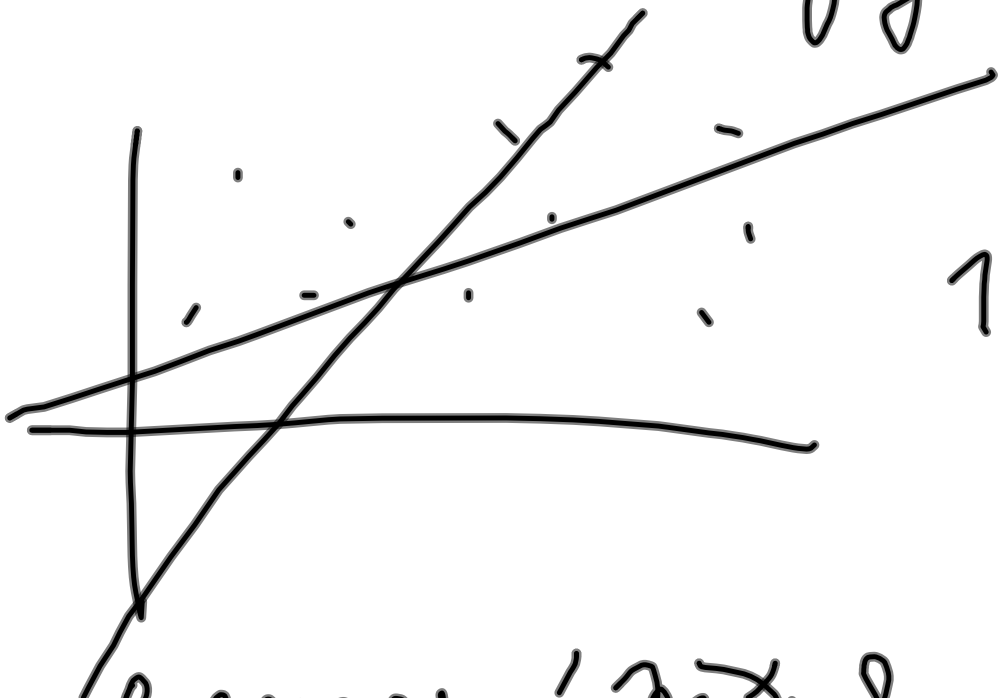
$$y-\bar{y}=\frac{k}{s_x^2}(x-\bar{x})$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad i \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$k = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

$x = \gamma y + \delta$



$$y - \bar{y} = \frac{k}{s_x^2}(x - \bar{x})$$

1. regresní přímka

2. regresní přímka

$$x - \bar{x} = \frac{k}{s_y^2}(y - \bar{y})$$

Když označíme $r = \dfrac{k}{s_x \cdot s_y}$ korelační koeficient

1. regresní přímka:

$$y - \bar{y} = r \cdot \frac{s_y}{s_x}(x - \bar{x})$$

2. regresní přímka

$$y - \bar{y} = \frac{1}{r} \cdot \frac{s_y}{s_x}(x - \bar{x})$$

Je-li $|r| = 1$, je $y = kx + q$

(~~ob~~ regresní přímky jsou totožné)

Hodnota $|r| < 1$ vyjadřuje míru lineární závislosti mezi x, y.