Odhad parametrů pro normální rozdělení

- parametry μ, σ $\quad (m = 2)$
- posteriorní hustota pravděpodobnosti: $f(\mu,\sigma|D_N) \propto P(D_N|\mu,\sigma) f(\mu,\sigma)$
- apriorní hustota pravděpodobnosti: $f(\mu,\sigma) \propto \text{konst.}$

$$f(\mu,\sigma|D_N) = \frac{1}{(2\pi)^{\frac{N}{2}}} \frac{1}{\sigma^N} \exp\left[-\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\ln f = -\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2} - N\ln(\sigma) - N\ln\left(\sqrt{2\pi}\right)$$

$$\left.\frac{\partial \ln(L)}{\partial \mu}\right|_{\mu_0,\sigma_0} = 0 \longrightarrow \sum_{i=1}^{N} \frac{x_i-\mu_0}{\sigma_0^2} = 0 \longrightarrow \boxed{\mu_0 = \frac{1}{N}\sum_{i=1}^{N} x_i \equiv \bar{x}}$$

$$\left.\frac{\partial \ln(L)}{\partial \sigma}\right|_{\mu_0,\sigma_0} = 0 \longrightarrow \sum_{i=1}^{N} \frac{(x_i-\mu_0)^2}{\sigma_0^3} - \frac{N}{\sigma_0} = 0 \longrightarrow \boxed{\sigma_0^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i-\bar{x})^2}$$

Přiřazení apriorní pravděpodobnosti

- princip invariance (Keynes 1921)

- **parametry polohy** – invariance vůči posunutí

$$f(x)\mathrm{d}x = f(x+x_0)\mathrm{d}(x+x_0) \longrightarrow f(x) = \text{konst.}$$

- **škálovací parametry** – invariance změně jednotek

$$f(y)\mathrm{d}y = f(\beta y)\mathrm{d}(\beta y) \longrightarrow f(y) \propto \frac{1}{y}$$

$$z \equiv \ln y \longrightarrow g(z) = \text{konst.}$$

Odhad parametrů pro normální rozdělení

- parametry μ, σ $\quad (m = 2)$
- posteriorní hustota pravděpodobnosti: $f(\mu,\sigma|D_N) \propto P(D_N|\mu,\sigma) f(\mu,\sigma)$
- apriorní hustota pravděpodobnosti: $f(\mu,\sigma) \propto \dfrac{1}{\sigma}$

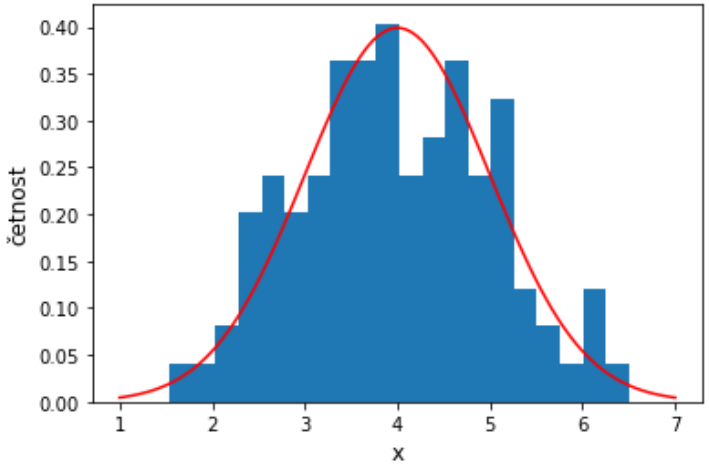
$$f(\mu,\sigma|D_N) = \frac{1}{(2\pi)^{\frac{N}{2}}} \frac{1}{\sigma^{N+1}} \exp\left[-\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\ln f = -\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2} - (N+1)\ln(\sigma) - N\ln\left(\sqrt{2\pi}\right)$$

$$\left.\frac{\partial \ln(L)}{\partial \mu}\right|_{\mu_0,\sigma_0} = 0 \longrightarrow \sum_{i=1}^{N} \frac{x_i-\mu_0}{\sigma_0^2} = 0 \longrightarrow \boxed{\mu_0 = \frac{1}{N}\sum_{i=1}^{N} x_i \equiv \bar{x}}$$

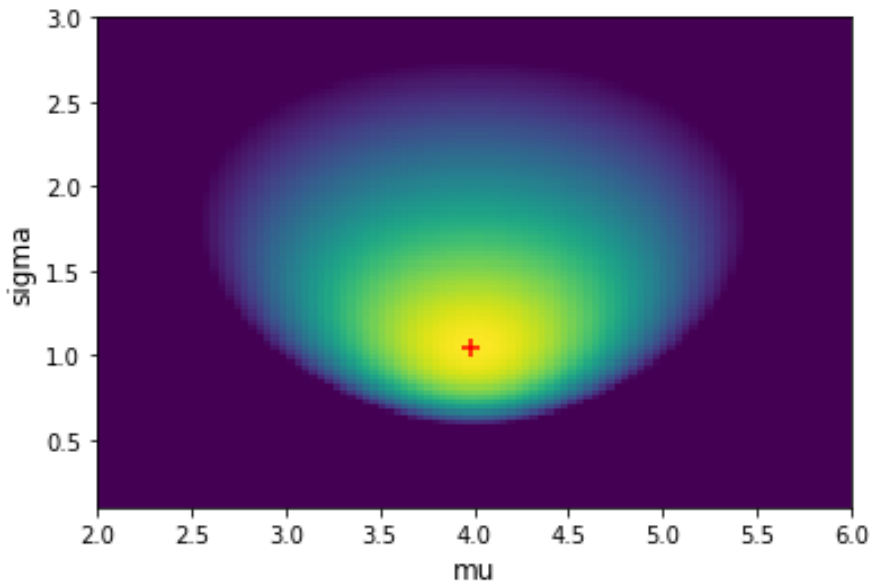
$$\left.\frac{\partial \ln(L)}{\partial \sigma}\right|_{\mu_0,\sigma_0} = 0 \longrightarrow \sum_{i=1}^{N} \frac{(x_i-\mu_0)^2}{\sigma_0^3} - \frac{N+1}{\sigma_0} = 0 \longrightarrow \boxed{\sigma_0^2 = \frac{1}{N+1}\sum_{i=1}^{N} (x_i-x)^2}$$

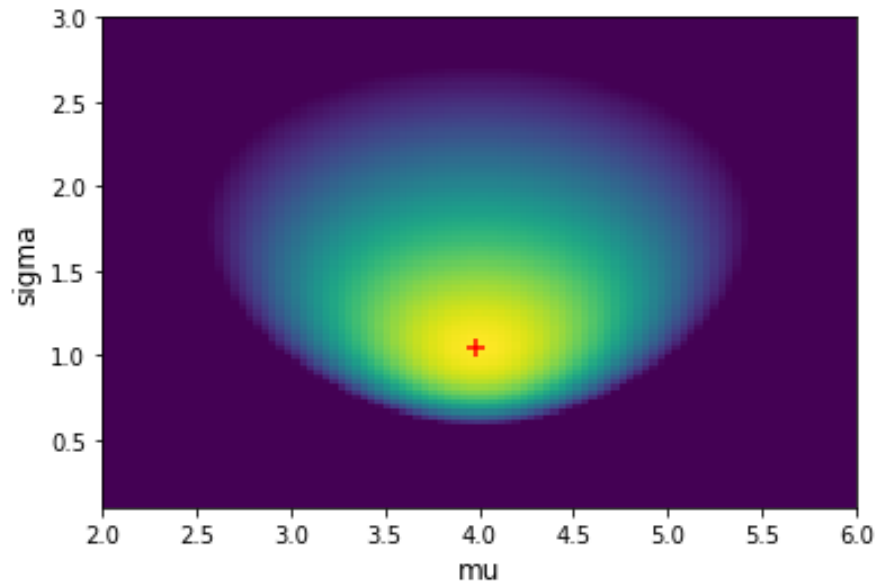
Odhad parametrů pro normální rozdělení



$$f(\mu,\sigma|D_N)=\frac{1}{(2\pi)^{\frac{N}{2}}}\frac{1}{\sigma^N}\exp\left[-\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$f(\mu,\sigma|D_N)=\frac{1}{(2\pi)^{\frac{N}{2}}}\frac{1}{\sigma^{N+1}}\exp\left[-\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$





Entropie

- entropie (Shannon 1948)

$$S = -\sum_{i=1}^{M} p_i \ln p_i$$

$$S = -\int p(x) \ln p(x) \, \mathrm{d}x$$

- **princip maximální entropie:**
  jako apriorní rozdělení bereme rozdělení s maximální entropií

Problém klokanů

- Gull & Skilling 1980

- 1/3 klokanů jsou leváci
- 1/3 klokanů mají modré oči
- jaké je procento klokanů leváků s modrýma očima?

		levák: ano	levák: ne
modré oči	ano	p_1	p_2
	ne	p_3	p_4

$$p_1 + p_2 + p_3 + p_4 = 1 \quad \text{(normalizace)}$$

$$p_1 + p_3 = \frac{1}{3} \quad \text{(leváci)}$$

$$p_1 + p_2 = \frac{1}{3} \quad \text{(modré oči)}$$

Problém klokanů

- Gull & Skilling 1980

- 1/3 klokanů jsou leváci
- 1/3 klokanů mají modré oči
- jaké je procento klokanů leváků s modrýma očima?

		levák: ano	levák: ne
modré oči	ano	$0 \le x \le \frac{1}{3}$	$\frac{1}{3} - x$
	ne	$\frac{1}{3} - x$	$\frac{1}{3} + x$

$$p_1 + p_2 + p_3 + p_4 = 1 \quad \text{(normalizace)}$$

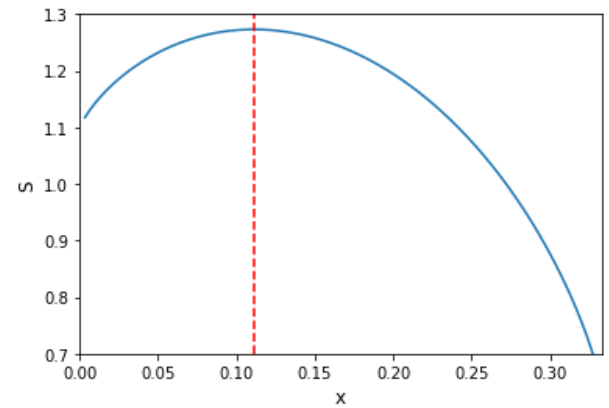
$$p_1 + p_3 = \frac{1}{3} \quad \text{(leváci)}$$

$$p_1 + p_2 = \frac{1}{3} \quad \text{(modré oči)}$$

$$x = \frac{1}{9}$$

Problém klokanů

- Gull & Skilling 1980



- 1/3 klokanů jsou leváci
- 1/3 klokanů mají modré oči
- jaké je procento klokanů leváků s modrýma očima?

		levák: ano	levák: ne
modré oči	ano	$\frac{1}{9}$	$\frac{2}{9}$
	ne	$\frac{2}{9}$	$\frac{4}{9}$

$$p_1 + p_2 + p_3 + p_4 = 1 \qquad \text{(normalizace)}$$

$$p_1 + p_3 = \frac{1}{3} \qquad \text{(leváci)}$$

$$p_1 + p_2 = \frac{1}{3} \qquad \text{(modré oči)}$$

$$x = \frac{1}{9}$$

- entropie $S = -\sum_{i=1}^{4} p_i \ln p_i = -x\ln(x) - 2\left(\frac{1}{3} - x\right)\ln\left(\frac{1}{3} - x\right) - \left(\frac{1}{3} + x\right)\ln\left(\frac{1}{3} + x\right)$

- maximální S $\longrightarrow$ $x = \frac{1}{9}$

Entropie (opičí tým)

- M možných výsledků ($x_1, x_2, \ldots x_M$)
- jak přiřadit pravděpodobnosti jednotlivým výsledkům?
- každou možnost reprezentujeme krabicí a náhodně do krabic rozházíme N mincí
- pravděpodobnost i-tého výsledku: $p_i = \frac{n_i}{N}$ (n_i – počet v mincí v i-té krabici)
- výsledkem je M-tice pravděpodobností: ($p_1, p_2, \ldots p_M$) $\sum_{i=1}^{M} p_i = 1$
- pokud to zopakujeme dostaneme jinou M-tici pravděpodobností
- frekvence výskytu M-tice ($p_1, p_2, \ldots p_M$): $F(\{p_i\}) = \frac{N!}{n_1! n_2! \ldots n_M!} M^{-N}$

$$\ln F = -N \ln M + \ln N! - \sum_{i=1}^{M} \ln n_i!$$

(Stirlingův vzorec: $\ln n! \approx n \ln n - n$)

$$\ln F = -N \ln M - N \sum_{i=1}^{M} p_i \ln p_i = \text{konst.} + NS$$

Entropie (opičí tým)

- M možných výsledků ($x_1, x_2, \ldots x_M$)
- jak přiřadit pravděpodobnosti jednotlivým výsledkům?
- každou možnost reprezentujeme krabicí a náhodně do krabic rozházíme N mincí
- pravděpodobnost i-tého výsledku: $p_i = \dfrac{n_i}{N}$ (n_i – počet v mincí v i-té krabici)
- pokud víme, že jednotlivé možnosti nejsou stejně pravděpodobné, zvolíme různě velké krabice
- pravděpodobnost, že mince padne do i-té krabice: m_i $\sum_{i=1}^{M} m_i = 1$
- frekvence výskytu M-tice ($p_1, p_2, \ldots p_M$): $F(\{p_i\}) = \dfrac{N!}{n_1! n_2! \ldots n_M!} m_1^{n_1} m_2^{n_2} \ldots m_M^{n_M}$

(multinomické rozdělení)

$$\ln F = \ln N! - \sum_{i=1}^{M} \ln n_i! + \sum_{i=1}^{M} n_i \ln m_i$$

(Stirlingův vzorec: $\ln n! \approx n \ln n - n$)

$$\ln F = -N \sum_{i=1}^{M} p_i \ln \frac{p_i}{m_i} = NS$$

Princip maximální entropie

- entropie (Shannon 1948)

$$S = -\sum_{i=1}^{M} p_i \ln p_i$$

$$S = -\int p(x) \ln p(x) \, \mathrm{d}x$$

- zobecněná entropie (Jaynes 1963)

$$S = -\sum_{i=1}^{M} p_i \ln \frac{p_i}{m_i}$$

$$S = -\int p(x) \ln \frac{p(x)}{m(x)} \, \mathrm{d}x$$

- **princip maximální entropie:**
  jako apriorní rozdělení bereme rozdělení s maximální entropií

m(x) Lebesqueova míra

zaručuje invarianci entropie při transformaci

$$x \to y = h(x)$$