# Estimating the fractal dimension with Takens' estimator

- Correlation dimension $\underset{\text{attractor}}{D_c(A)} = \lim_{\varepsilon \to 0} \frac{\ln P(\varepsilon)}{\ln \varepsilon}$ ← $P(\varepsilon) = \underset{\text{product measure}}{m^2}(\{(a_1,a_2) \in A \times A : \underset{\text{euclidean distance}}{\delta(a_1,a_2)} < \varepsilon\})$

- Estimation of the correlation dimension: Take finite sequence $\delta_1, \dots, \delta_m$ of distances between randomly chosen pairs $(a_1,a_2)$. If $D_c$ exists, then $P(\varepsilon) = C \cdot \varepsilon^\alpha$ + higher order terms.

→ **Assumption**: $\exists\, \varepsilon_0 > 0$ fixed, s.th. $\forall\, 0 \le \varepsilon \le \varepsilon_0$ $P(\varepsilon) = C \cdot \varepsilon^\alpha$ exactly.

↳ Now take all $\varepsilon < \varepsilon_0$, normalize s.th. $r_i = \frac{\delta_i}{\varepsilon_0} < 1$

↳ Interpret $P(\varepsilon)$ as probability: distribution of $r_i$ is given by

$$P(0 \le r_i \le t) = t^\alpha \quad \text{for } t \in [0,1]$$

$\Rightarrow$ $y_i = -\ln r_i$ is distributed exponentially with parameter $\alpha$.

Proof: $X$ is RV with cdf

$$P(0 \le X \le t) = \underbrace{t^\alpha}_{\text{cdf}} = \int_0^t \underbrace{\alpha s^{\alpha-1}}_{\text{pdf}} ds$$

$\Rightarrow$ for $Y = f(X) = -\ln X$,

$$P(Y \le y) = P(-\ln X \le y) = P(X \ge e^{-y}) = \int_{e^{-y}}^1 \alpha s^{\alpha-1} ds = s^\alpha \Big|_{e^{-y}}^1 = \underbrace{1 - e^{-\alpha y}}_{\text{cdf of exp. dist.}}$$

**MLE** for exp. dist.:

likelihood-function* ↘ $L(\alpha, \{\ \}) = \prod_{j=1}^m \mathrm{pdf}(y_j; \alpha) = \prod_{j=1}^m \alpha \cdot \exp(-\alpha y_j)$

$\Rightarrow$ $\ell(\alpha, \{y_i\}) = \ln L(\alpha, \{y_i\}) = m \ln \alpha - \alpha \sum_{i=1}^m y_i$

→ Question: For which parameter $\alpha$ is likelihood of obtaining our sample $\{y_i\}$ maximal?

$$\hat{\alpha}_{MLE} = \underset{\alpha}{\operatorname{argmax}}\ \ell(\alpha; \{y_i\})$$

$$\Rightarrow 0 \overset{!}{=} \frac{d}{d\alpha} \ell(\alpha; \{y_i\}) \Big|_{\alpha = \hat{\alpha}_{MLE}} = \frac{m}{\hat{\alpha}_{MLE}} - \sum_{i=1}^m y_i$$

$\Rightarrow$ $\hat{\alpha}_{MLE} = \frac{m}{\sum_{i=1}^m y_i} = \frac{1}{\bar{Y}}$ ↖ sample mean (remember: $Y_j = -\ln r_j$)

This is actually biased:

$y_i$ are exp. dist, which is gamma dist w/ shape param. 1

↳ $\sum_{i=1}^m y_i \sim \mathrm{Gamma}(m, \alpha)$

↳ $\frac{1}{\sum_{i=1}^m y_i} \sim \mathrm{Inv.Gamma}(m, \alpha)$.

$$\langle \hat{\alpha}_{MLE} \rangle = \left\langle \frac{m}{\sum_{i=1}^m y_i} \right\rangle = m \left\langle \frac{1}{\sum_{i=1}^m y_i} \right\rangle = m \underbrace{\frac{\alpha}{m-1}}_{\text{expectation value of Inv.Gamma}(m,\alpha)} \neq \alpha$$

Correction: $\boxed{\hat{\alpha}^* = \frac{m-1}{m} \hat{\alpha}_{MLE} = \frac{m-1}{\sum_{i=1}^m y_i}}$

---

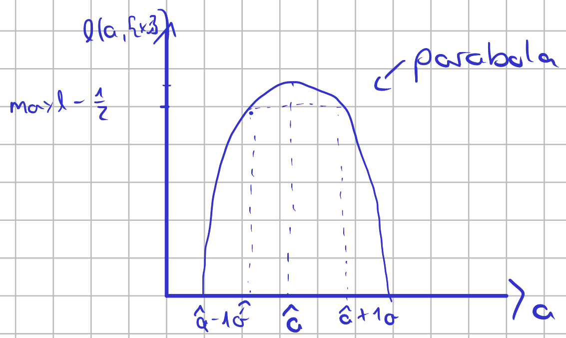
* likelihood of obtaining a sample of measurements $\{x_i\}$

# Estimation of confidence limits for MLE

- ML-estimators are invariant: $\widehat{f(a)} = f(\hat{a})$, where $\hat{a}$ is the maximum of $L$

$\Rightarrow$ For a Gaussian distr. variable, the log. likelihood function would be a parabola ($L(a, \{x_i\}) \propto \exp\left[A \cdot \frac{(a-\hat{a}(x_1, \ldots, x_n))^2}{2}\right]$). Standard deviation of $L$ is $1/\sqrt{A}$, which is also the standard deviation of $\hat{a}$ $\left(\sigma_a^2 = -\frac{1}{\left(\frac{d^2 \ln L}{da^2}\Big|_{a_0}\right)}\right)$

$\Rightarrow$ We can read the errors from the graph of $l$:




At $1\sigma$, the graph has fallen by $\frac{1}{2}$ from it's maximum, at $2\sigma$, it has fallen by $2$.

- If the CLT is not applicable (yet), we can estimate $\sigma$ using the invariance:

There is some $a' = f(a)$ for which $\ln L(a')$ is a parabola. For this, the $\pm\sigma$ limits are where the parabola has fallen by $\frac{1}{2}$ from its maximum. By invariance, the corresponding values for $a$ (where $\ln L(a)$ is not parabolic) are the values for which $\ln L(a)$ has fallen by $\frac{1}{2}$ from its maximum.

02.10.2020

# Fractal Dimension estimation with Takens' estimator: Pros & Cons

- Nice properties
  - ↳ We can actually estimate the confidence limits, even asymmetric ones
  - ↳ Computationally quite inexpensive
- Not-so-nice properties
  - ↳ Estimation + confidence limits are under the assumption that $C(\varepsilon) = \varepsilon^{\alpha}$ exactly, technically we would have to calculate $C(\varepsilon)$ for a lot of different $\varepsilon$ to find out the interval for which that is true (if it exists), and then we can really just use the grassberger method in the first place, especially since we get variations of like $10 \cdot 95\%$ conf-interval with different values for $\varepsilon_{max}$

  ⇒ Pretends to be really accurate (which it is if the assumption is true, but we don't really know when that's the case), introducing a false sense of high accuracy

⇒ Judgement: Interesting, fast, but ultimately useless (unless used on top of Grassberger for interval where we know $C(\varepsilon) = \varepsilon^{\alpha}$ holds to get proper confidence limits, but then it's not fast anymore)

So, the only advantage is the reliable estimation of confidence limits if one cares about that (under the assumption $C(\varepsilon) = \varepsilon^{\alpha}$!)

(Maybe the way that $D_c$ depends on $\varepsilon_{max}$ could be interesting?)