Classical statistical inference

Uncertainties on arbitrary RV

Part 4

Associated notebook:

04-Basic_statistical_inference_frequentists_2/Frequentist_Monte_Carlo.ipynb

Uncertainty calculation

How to calculate stde (or simply std) on a RV if it is a function of one or multiple RV ?

Case 1: $y = \phi(x)$ $\quad$ $p(x)$ known

Case 2: $z = \phi(x, y)$ $\quad$ $p(x)$ and $p(y)$ known

Uncertainty calculation

Case 1: $y = \phi(x)$ $p(x)$ known

Example: $y = -2.5 * log(x)$

Uncertainty calculation

Case 1: $y = \phi(x)$ $p(x)$ known

Example: $y = -2.5 * log(x)$

Variable transformation: $p(y) = P'(y) = p\left[\phi^{-1}(y)\right]\left|\frac{\mathrm{d}\,\phi^{-1}(y)}{\mathrm{d}y}\right|$

Uncertainty calculation

Case 1: $y = \phi(x)$ $p(x)$ known

Example: $y = -2.5 * log(x)$

Variable transformation: $p(y) = P'(y) = p\left[\phi^{-1}(y)\right] \left| \frac{\mathrm{d}\,\phi^{-1}(y)}{\mathrm{d}y} \right|$

$x = \phi^{-1}(y) = 10^{-0.4*y}$

$p(x) \sim U(0,1)$

1 if x ∈ [0,1]

0 if x ∉ [0,1]

$p(y) = 1 \times 0.4\,\ln(10)\,10^{-0.4\,y}$

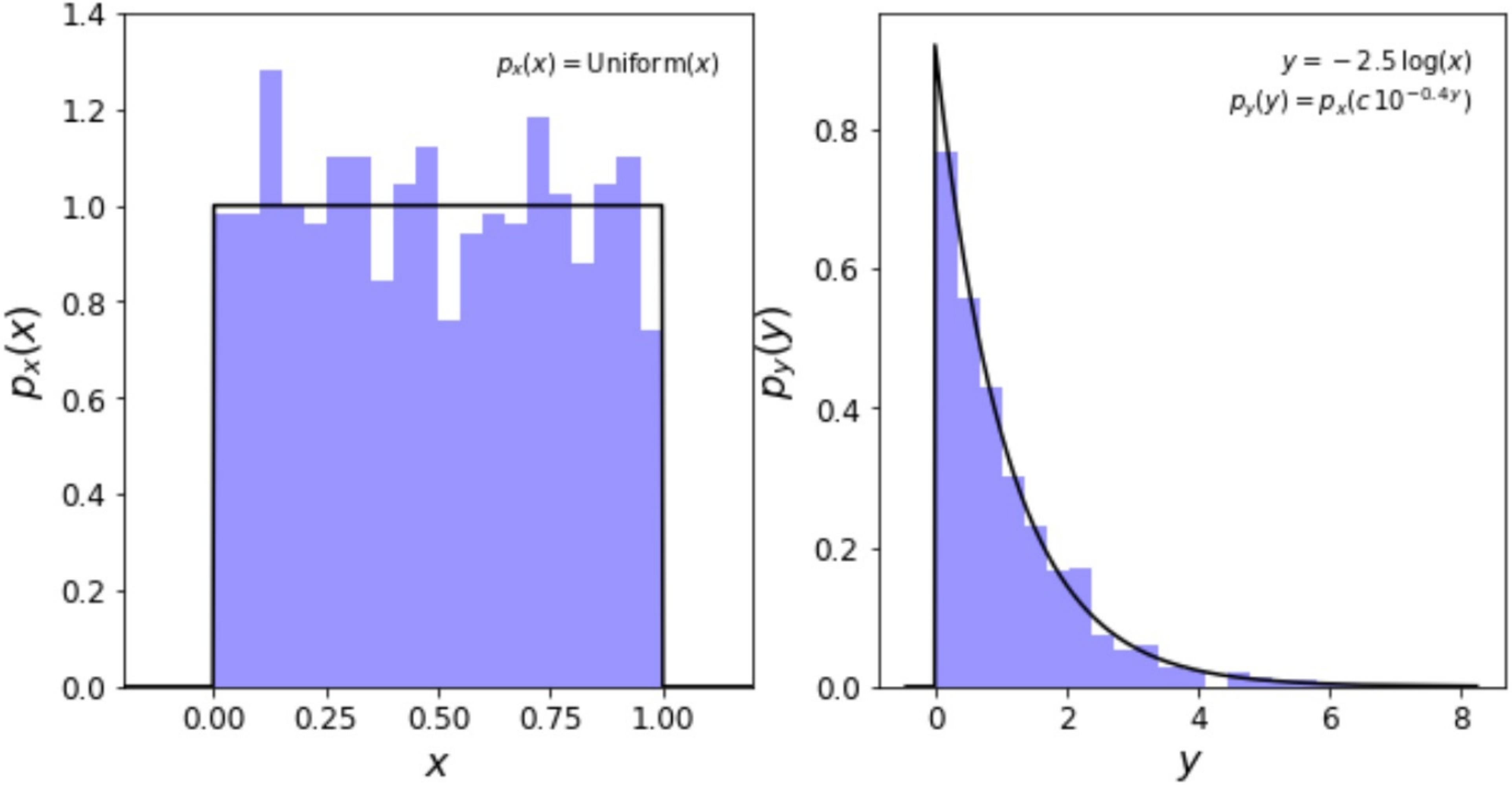
$0 < y < \infty$

Uncertainty calculation

Go to: Sect. II.4.1 of the notebook

See also II.4.1.1 and meaning of dex

Example: $y = -2.5 * log(x)$



Uncertainty calculation

Case 2: $z = \phi(x, y)$ $p(x)$ and $p(y)$ known

Error propagation formula

$$\sigma_z^2 = \left(\frac{\partial \phi}{\partial x}\right)_{\bar{x}}^2 \sigma_x^2 + \left(\frac{\partial \phi}{\partial y}\right)_{\bar{y}}^2 \sigma_y^2 + 2\left(\frac{\partial \phi}{\partial x}\right)_{\bar{x}} \left(\frac{\partial \phi}{\partial y}\right)_{\bar{y}} \sigma_{xy}$$

Results from Taylor expanding *z* around $\bar{x}$ and $\bar{y}$ => neglects some high order terms

Go to: Sect. II.4.2 of the notebook for the demo

Uncertainty calculation *Monte-Carlo* $c = \Phi(a, b)$: σ_c ?

The expectation ("average") of a function *f(x)* of a RV *x* can be approximated by drawing a virtually infinite sample from *x*

$$E(f(x)) = \int_{-\infty}^{+\infty} f(x)h(x)dx \rightarrow \frac{1}{N}\sum_{i}^{N} f(x_i),$$

You can specialise *f() to the calculation of* the mean, or the variance of a RV.

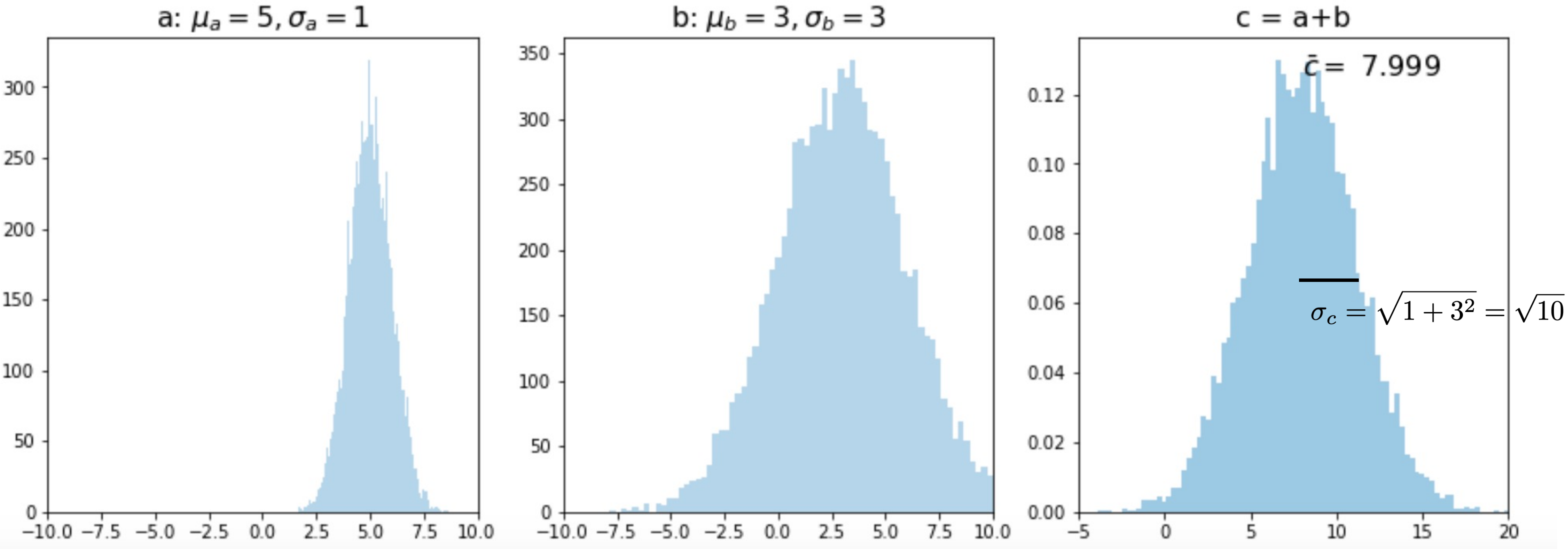
1. Draw random samples from $h(a; \mu_a, \sigma_a)$, $h(b, ; \mu_b, \sigma_b)$
2. Construct a random (*monte-carlo*) sample of $c_i = \Phi(a_i, b_i)$
3. Derive σ_c from your Monte-Carlo sample

Uncertainty calculation *Monte-Carlo*

Go to: Sect. II.4.3 of the notebook

c = a + b

1. Draw random samples from $h(a;\, \mu_a,\, \sigma_a)$, $h(b,\, ;\, \mu_b,\, \sigma_b)$
2. Construct a random (*monte-carlo*) sample of $c_i = \Phi(a_i, b_i)$
3. Derive σ_c from your Monte-Carlo sample



Classical statistical inference

Bootstrap and Jacknife

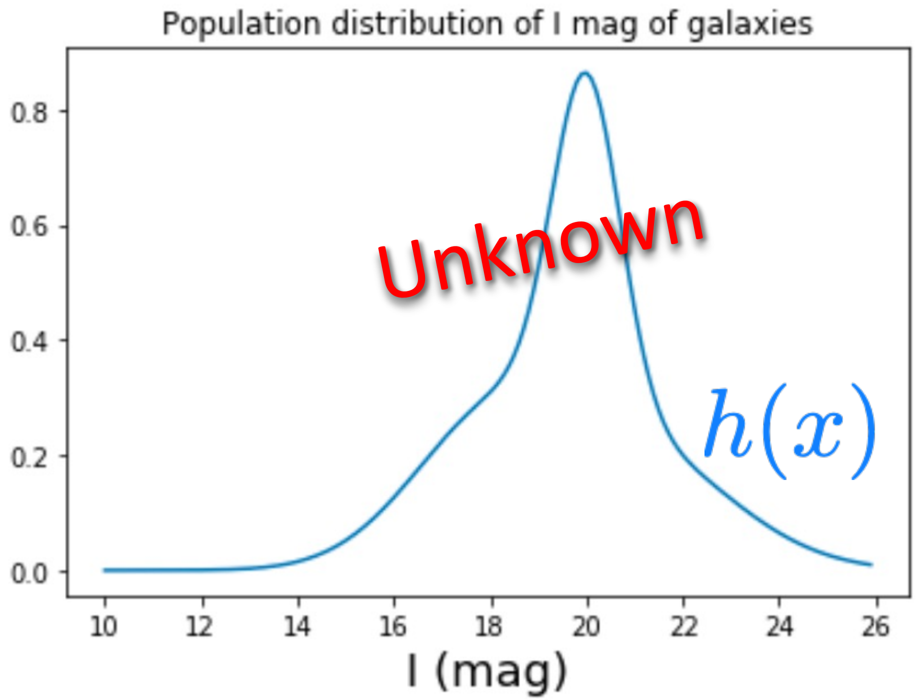
Part 5

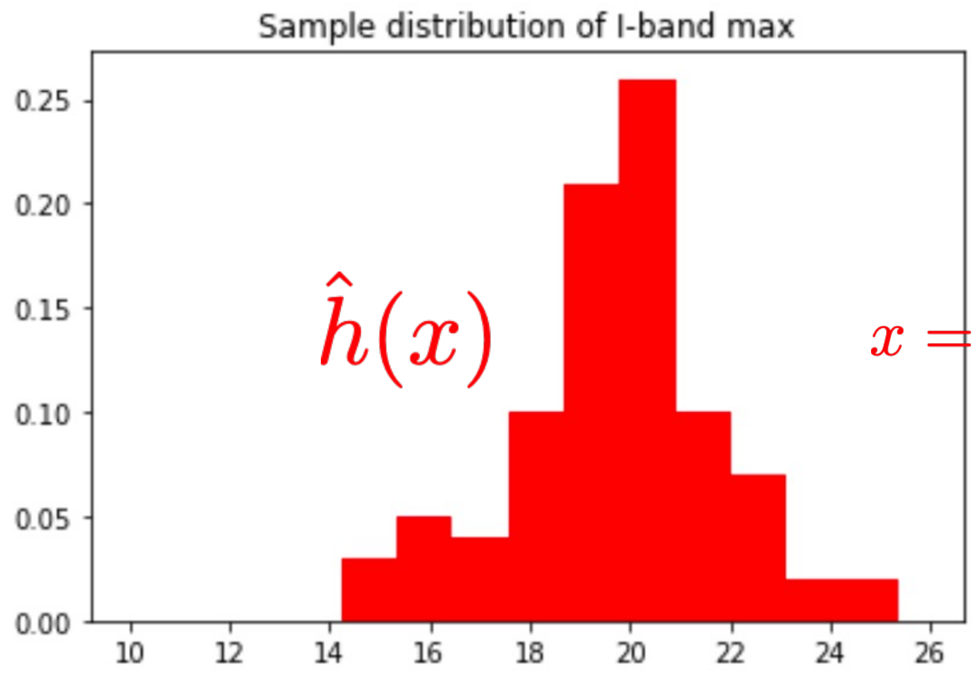
Associated notebook:

[04-Basic_statistical_inference_frequentists_2/
Frequentist_inference_Bootstrap.ipynb]

Bootstrap

Go to: Sect. II.5 of the notebook





My data:

$$x = \{x_1, x_2, ..., x_n\}$$

Bootstrap ≡ Draw samples from the sample PDF $\hat{h}(x)$, allowing from replacements

$$B = \{x_1^*, x_2^*, ..., x_n^*\}$$

$$\text{w.}\, x^* \text{ from } \{x_1, x_2, ..., x_n\}$$

e.g.

$$B1 = \{x_1, x_1, x_7, ..., x_{28}\}$$

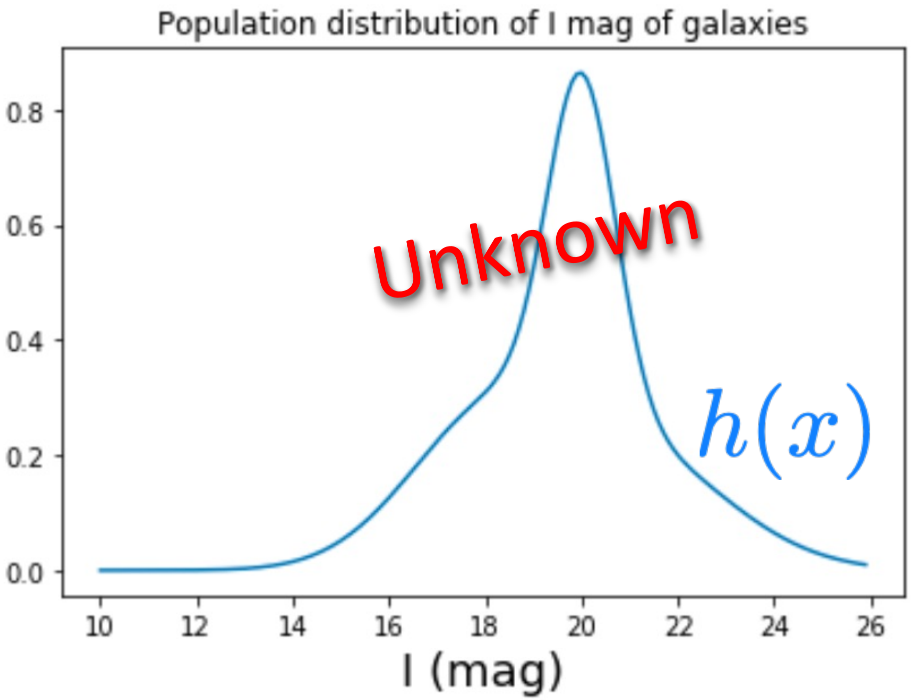
$$B2 = \{x_2, x_{36}, x_9, ..., x_8\}$$

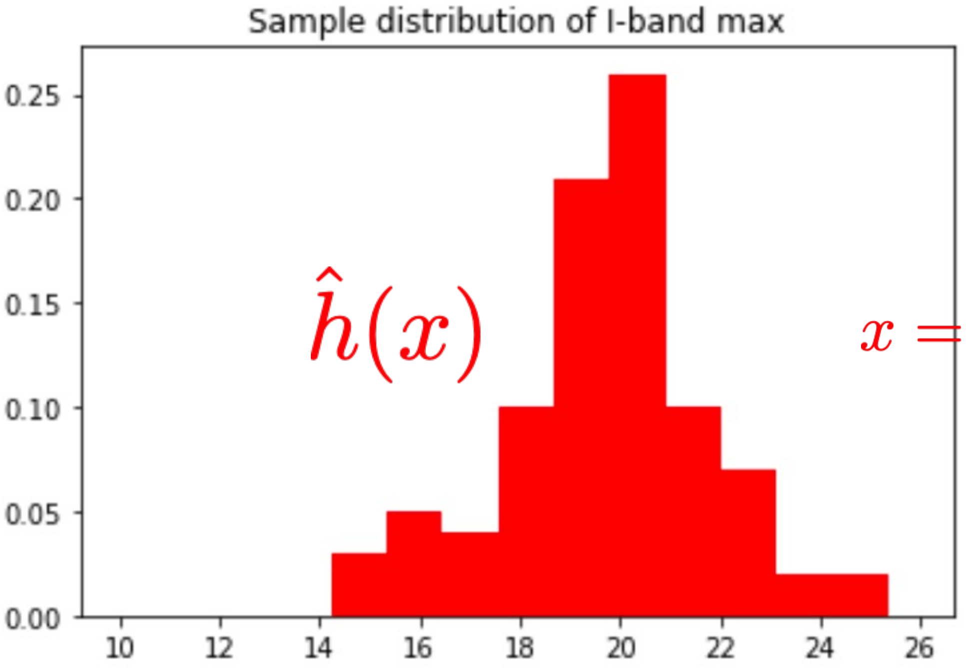
...

$$Bk = \{x_{16}, x_{12}, x_3, ..., x_{10}\}$$

Bootstrap Confidence Interval

Go to: Sect. II.5.1 of the notebook





My data:

$$x = \{x_1, x_2, ..., x_n\}$$

Calculate the estimate of your statistics *q* from all the bootstrapped samples + its associated $stde^B(q)$

Normal CI:

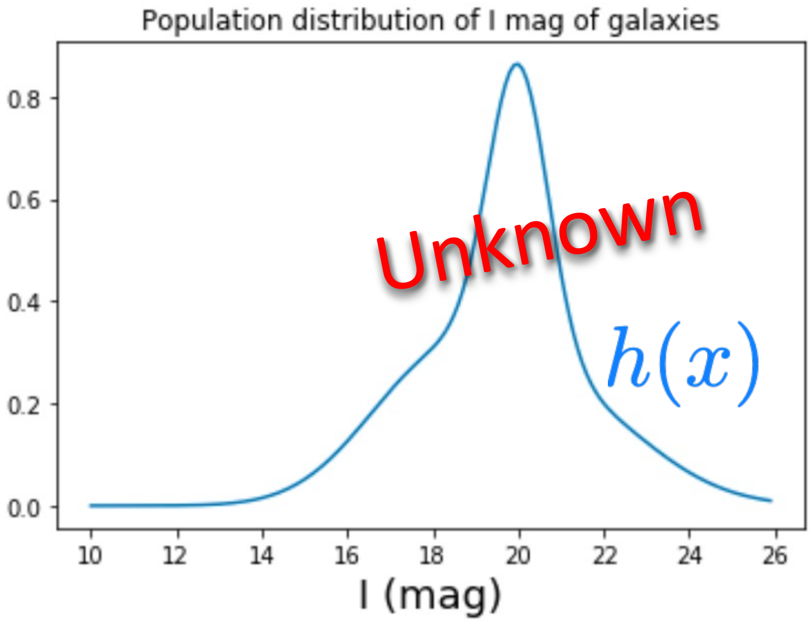
$$[\hat{q} - z_{\alpha/2}\, stde^B(q), \hat{q} + z_{\alpha 2}\, stde^B(q)]$$

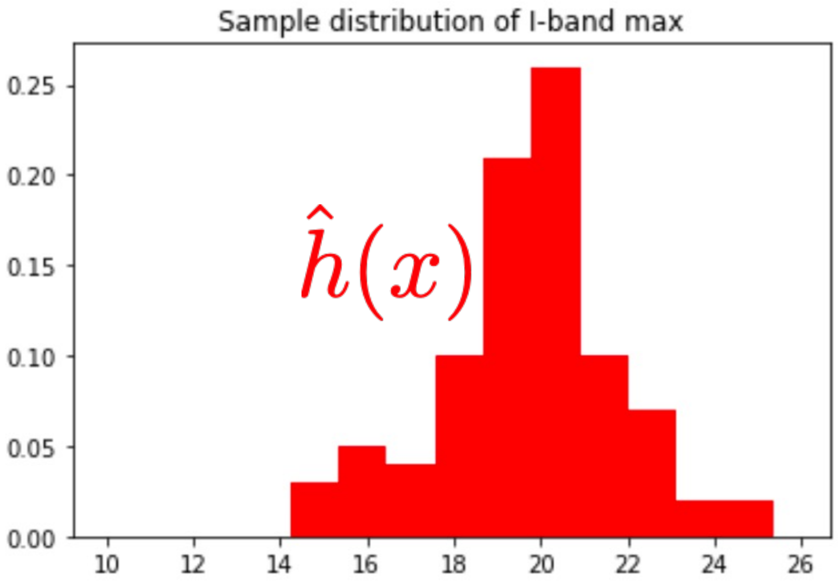
Percentile CI:

$$[q^*_{\alpha/2}, q^*_{1-\alpha/2}]$$

Jacknife

Go to: Sect. II.5.2 of the notebook





My data:

$$x = \{x_1, x_2, \ldots, x_n\}$$

$$J1 = \{x_1, \ldots, x_{n-2}, x_{n-1}\}$$

$$J2 = \{x_1, \ldots, x_{n-2}, x_n\}$$

$$\ldots$$

$$Jn = \{x_2, \ldots, x_{n-1}, x_n\}$$

With n-1 point per sample

Jacknife ≡ Remove *1* data point from your sample, *n* times

WARNING: statistics q_n from Jacknife is biased

$$q^J = n\, q_n - (n-1)\, \bar{q}_n$$

w. $$\bar{q}_n = n^{-1} \sum_{i=1}^{n} q_i^*$$

$$\sigma_q^2 = \frac{n-1}{n} \sum_{i=1}^{n} (q_i^* - \bar{q}_n)^2$$