PDF estimation by use of characteristic functions and the FFT

Jont B. Allen

Univ. of IL,

Beckman Inst., Urbana IL

Given N i.i.d. samples of a r.v., find the PDF

- The sample PDF

$$\widetilde{p}_x(\zeta, N) = \frac{1}{N}\sum_{n=1}^{N}\delta(\zeta - \tilde{x}_n).$$

- The Fourier transform gives the *sample characteristic function* (CF)

$$\widetilde{P}_x(\nu, N) = \frac{1}{N}\sum_{n=1}^{N} e^{-\jmath\nu\tilde{x}_n}.$$

Given N i.i.d. samples of a r.v., find the PDF

- For a two–dimensional r.v. $(\tilde{x}_n, \tilde{y}_n)$

$$\widetilde{p}_{x,y}(\zeta_1, \zeta_2, N) = \frac{1}{N} \sum_{n=1}^{N} \delta(\zeta_1 - \tilde{x}_n)\, \delta(\zeta_2 - \tilde{y}_n),$$

  giving a *sample 2-D characteristic function*

$$\widetilde{P}_{x,y}(\nu_1, \nu_2, N) = \frac{1}{N} \sum_{n=1}^{N} e^{-\jmath(\nu_1 \tilde{x}_n + \nu_2 \tilde{y}_n)}.$$

- . . . and so on . . . for any number of dimensions

Error I: Sampling uncertainty

- The *sampling uncertainty*, due to the finite sample size, is defined as

$$\widetilde{U}(\nu, N) = \widetilde{P}_x(\nu, N) - P_x(\nu),$$

  - The expected value of $\widetilde{P}_x(\nu, N)$ is $P_x(\nu)$ (i.e., $\mathcal{E}[\widetilde{U}(\nu, N)] = 0$)
  - The variance of $\widetilde{U}(\nu, N)$ may be shown to be:

$$\begin{aligned} \sigma^2_{\mathsf{U}}(\nu) &= \mathcal{E}[|\widetilde{U}(\nu, N)|^2] \\ &= \frac{1 - |P_x(\nu, N)|^2}{N} \\ &< 1/N \end{aligned}$$

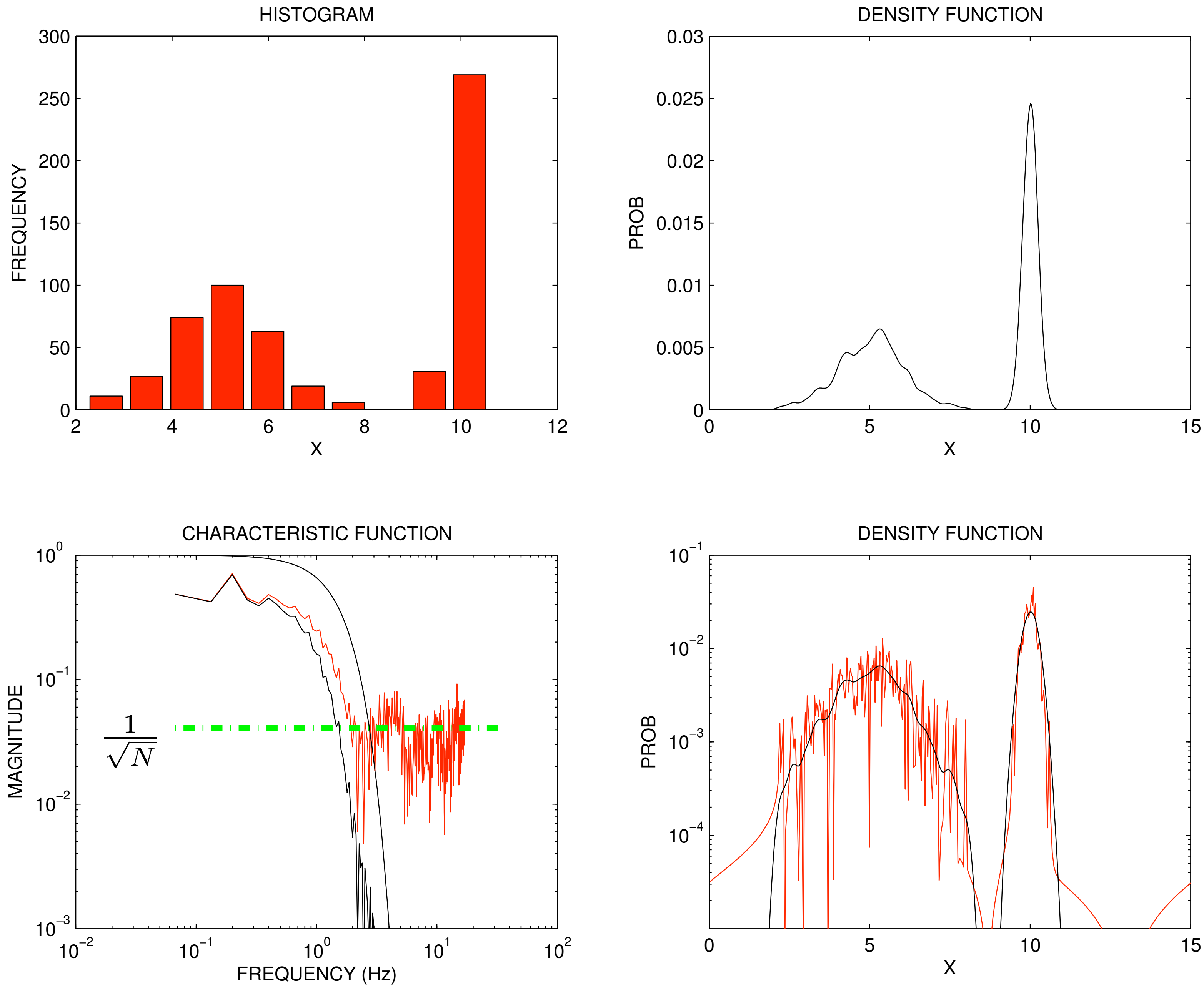
- IID Sampling noise is close to "*white*."

Smoothing

- The next step is to form the estimate $\widehat{P}_x(\nu)$ by filtering $\widetilde{P}_x(\nu)$ with a *data dependent filter*.
- The filter design is a classical Wiener–filtering problem
- The CF–domain Wiener filter may be designed based on the precise knowledge of the variance of the sampling uncertainty, i.e.

$$\sigma^2_{\mathsf{U}}(N) = \frac{1 - |P(\nu, N)|^2}{N} < \frac{1}{N}$$

Example: Wiener filter in the CF domain



A low-pass filter in the CF domain

- Low pass filtering in the CF domain allow for sampling in the probability domain. (Nyquist samples)
- Thus after this LP Filter, we may down-sample (quantize) the data
- Quantizing the data leads to alias images
  - i.e., from the *Poisson Sampling formula*
- Nyquist Sampling PDF levels $\approx$ quantizing levels

Error II: Binning errors

- Quantizing (binning) error may be computed from

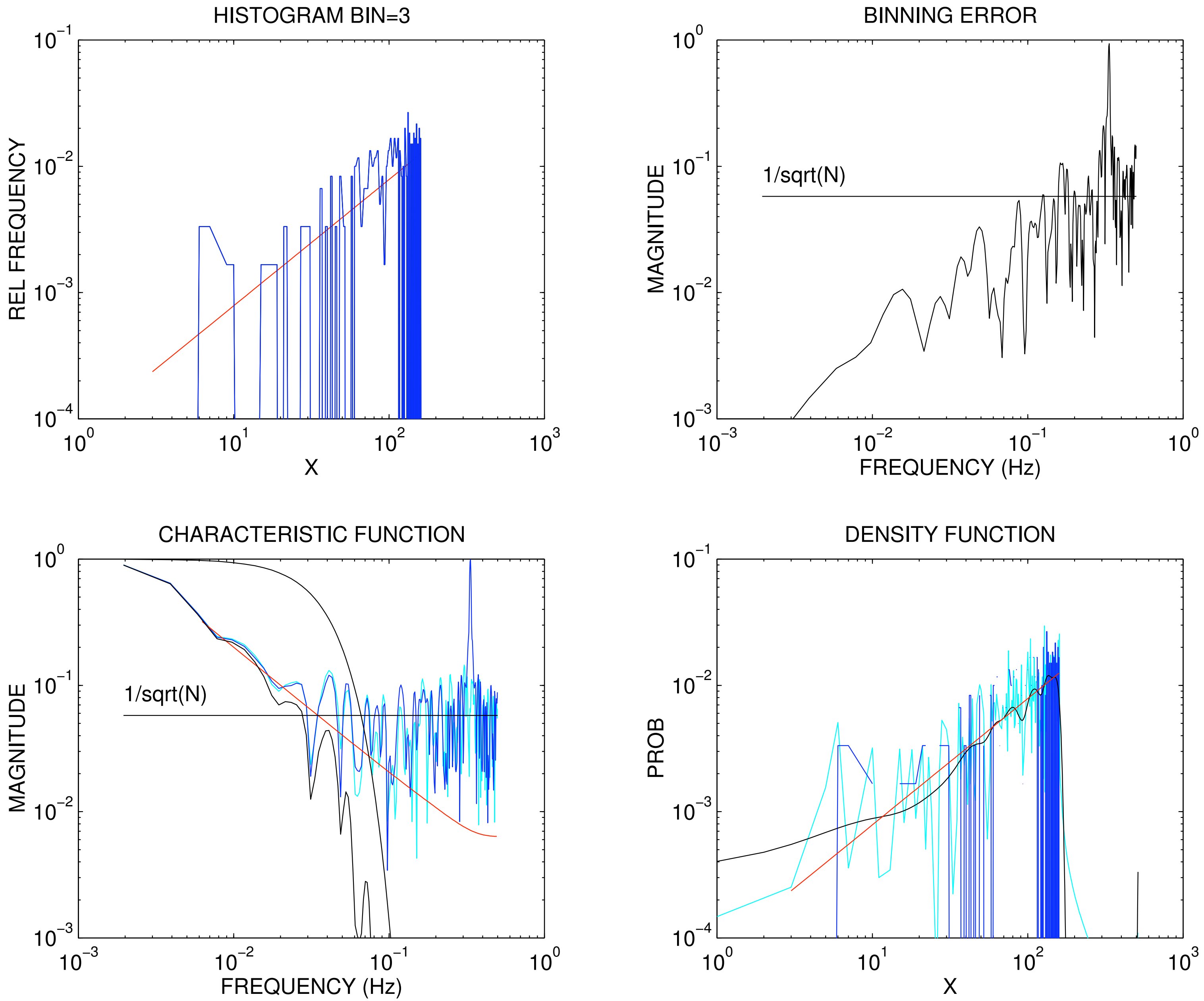
$$\widetilde{Q}(\nu, N) = \frac{1}{N}\sum_{n} e^{-i\nu\lfloor\tilde{x}_n\rceil} - \widetilde{P}_x(\nu, N),$$

- Example: 300 samples were drawn from

$$\begin{aligned} p_y(\zeta, 300) &= \zeta/12775, \quad for\, 2.26 \leq \zeta \leq 159.86 \\ &= 0, \quad otherwise, \end{aligned}$$

and were quantized with a bin width of 2

Binning errors (Error type II)



Alternative to Wiener Filter

- The Wiener Filter is a bit clumsy
  - I.E. what phase for the filter should we take?
  - How to best pick the shape of the filter?
- An alternative would be to separate the two distributions with the EM algorithm
- One PDF is the desired one, the second is the sampling noise.
  - Use parametric models of the two PDFs.
  - Use the near–white property of the sample noise.
- I believe this has great merit, and seems to be a novel approach

NEW TOPIC

- Nyquist Sampling and Characteristic Functions
- What does it mean to have a “band-limited” CF?
- If the CF is “band-limited” then it may be sampled
- Such a sampling corresponds to level quantization
  - This leads to *aliasing* of the CF

The Nyquist theorem

- When two independent random variables are added, we know that their PDF's are convolved.
- Thus in the CF domain:

$$P_{s+n}(\nu) = P_s(\nu)\, P_n(\nu)$$

- One interpretation of this relation is that:
  - The CF for the noise acts like a "lowpass filter" on the CF of the signal.
- If the CF is sufficiently "band limited," we may "sample" the PDF at the "Nyquist rate", e.g.

$$p_{s+n}(k\sigma_n/2), k = \cdots, -1, 0, 1, 2, \cdots$$

with *no loss of fidelity.*

Gaussian example

- Suppose that n is i.i.d. and $\mathcal{N}(0, \sigma_n)$. Find ζ_k, the Nyquist samples.
- The Fourier transform pair are

$$
\begin{aligned}
p_n(\zeta) &= \frac{e^{-\zeta^2/2\sigma_n^2}}{\sigma_n\sqrt{2\pi}} \\
P_n(\nu) &= e^{-\sigma_n^2\nu^2/2}.
\end{aligned}
$$

- Define the maximum radian frequency ν_{max} as

$$\nu_{max} = 2\pi/\sigma_n.$$

  - Then $P_n(\nu_{max}) = 2.68 \times 10^{-9}$, and the aliasing (i.e., the error due to sampling) will be negligible.

Gaussian example, cont.

- The Nyquist theorem says that if we know $p_n(\zeta)$ at $\zeta = \zeta_k$, where k is an integer and

$$\begin{aligned} \zeta_k &= k\,2\pi/2\nu_{max} \\ &= k\sigma_n/2, \end{aligned}$$

we may reconstruct $p_n(\zeta)$ from fixed samples with negligible error (e.g., an error of about $P_n(\nu_{max})$).

Testing for independence

- Suppose we wish to test the independence of variables x and y, from samples from P_x and P_y. To do this we test the ratio

$$\frac{P(x,y)}{P_x \, P_y}$$

to see if it is statistically different from one.

  - Example:

$$P(x,y) = P(s_n, a(s_n - s_{n-1})).$$

- To implement such a test, we must identify the frequency regions where the CF is greater than the sampling uncertainty noise floor, as given by $1/\sqrt{N}$. Then we look at how the ratio differs from 1 over this region. If it differs from 1 by more than $1/\sqrt{N}$, the two signals are not independent.

Example problems

- Question (Quantization): "If we have a signal x, and we send it over a channel, giving a new signal $y = x + n$, can we distinguish this new signal from one that has been quantized before it was sent over the same channel? In other words, can we distinguish the signals $x + n$ and $Q(x) + n$, where $Q(x)$ represents the uniform-level Nyquist quantizer?
- Question (Representation): Given a set of random variables, how many discrete levels are required to represent the values?
  - After sampling the PDF, we may count the number of *states*. Let $R_{max} = \max(s) - \min(s)$. Then

$$\mathcal{C} = \log_2 \left(1 + \nu_{max}\, R_{max}/\pi\right)$$