# Signals and Spectra (2)

Appendix B

Lecture 4, 2008-09-16

# Introduction

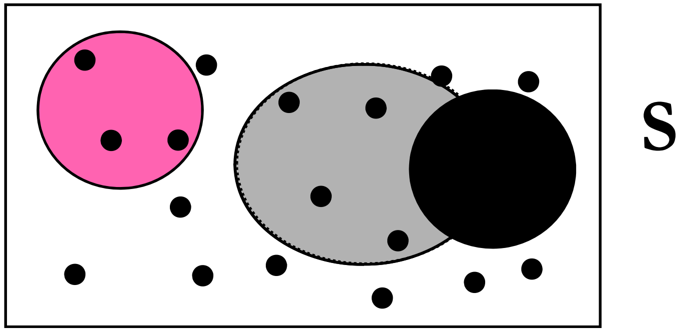
- Probability and Random Variables
- Cumulative Distribution Functions and Probability Density Functions
- Ensemble Average and Moments
- Functional Transformations of Random Variables
- Multi-Variable and Bi-Variable Statistics
- Central Limit Theorem

# Probability

- P(A), the probability of an event A, is defined in terms of the relative frequency of A occurring in n trails.

$$P(A) = \lim_{n \to \infty} \left( \frac{n_A}{n} \right)$$

- Mathematically characterizes random events.
- Defined on a probability space: (S,$\mathcal{E}$,P(•))
  - Sample space of possible outcomes.
- Sample space has a subset of events
- Probability defined for these subsets.



# The Axioms of Probability

- Axiom 1. $P(A) \geq 0$ for all events A in the sample space S.
- Axiom 2. The probability of all possible events occurring is unity, $P(S) = 1$.
- Axiom 3. If the occurrence of A precludes the occurrence of B, and vice vera (i.e. A and B are mutually exclusive), then $P(A \cup B) = P(A) + P(B)$

# Joint and Conditional Probability

- The probability of a joint event AB, is

$$P(AB) = \lim_{n \to \infty} \left( \frac{n_{AB}}{n} \right)$$

Where $n_{AB}$ is the number of times that the event AB occurs in n trials.

- Theorem: Let E = A + B; then P(E) = P(A) + P(B) – P(AB)
- The probability that an event A occurs, given that an event B has also occurred, is denoted by P(A|B), which is defined as

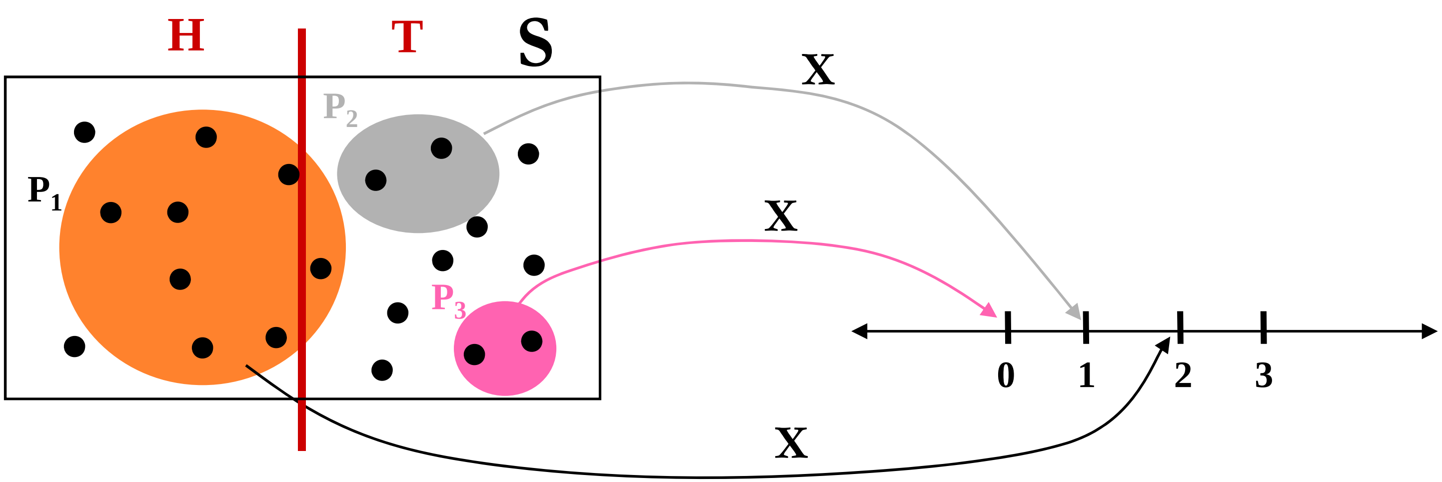
$$P(A \mid B) = \lim_{n_B \to \infty} \left( \frac{n_{AB}}{n_B} \right)$$

- Theorem: Let E = AB; then P(E) = P(A)P(B|A) = P(B)P(A|B)

  This is known as Bayes' Theorem.
- Two events, A and B, are said to be independent if either P(A|B) = P(A) or P(B|A) = P(B)

# Random Variables

- A real-value random variable is a real-valued function defined on the events of the probability system.
- In the applications of probability it is more convenient to work in terms of numerical outcomes (e.g. the number of errors in a digital data message) rather than non-numerical outcomes (e. g. failure of a component)



# Cumulative Distribution Function

- The cumulative distribution function (CDF) of the random variable x is given by

$$F(a) \equiv P(x \le a) \equiv \lim_{n\to\infty}\left(\frac{n_{x\le a}}{n}\right)$$

F(a) is a unitless function

- CDF has following properties:
  - 1. $0 \le F(a) \le 1$, with $F(-\infty) = 0$ and $F(\infty) = 1$
  - 2. F(a) is a non-decreasing function of x; that is $F(x1) \le F(x2)$ if $x1 \le x2$
  - 3. F(a) is continuous from the right $\lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} F(a+\varepsilon) = F(a)$

# Probability Density Function

- The probability density function (PDF) of the random variable x is given by

$$f(x) = \left.\frac{dF(a)}{da}\right|_{a=x} = \left.\frac{dP(x \le a)}{da}\right|_{a=x} = \lim_{\substack{n\to\infty \\ \Delta x\to 0}} \left[\frac{1}{\Delta x}\left(\frac{n_{\Delta x}}{n}\right)\right]$$

f(a) has units of 1/x

- PDF has following properties:
  - 1. $f(x) \ge 0$. That is, f(x) is a nonnegative function
  - 2. $\int_{-\infty}^{\infty} f(x)dx = F(+\infty) = 1$

# Ensemble Average and Moments

- The expected value, which is also called the ensemble average, of y = h(x) is given by $\bar{y} = \int_{-\infty}^{\infty} [h(x)] f(x) dx$

The ensemble average of y is also denoted by E[y] or<y>.

- The rth moment of the random variable x taken about the point x = x0 is given by $\overline{(x - x_0)^r} = \int_{-\infty}^{\infty} (x - x_0)^r f(x) dx$
- The mean is the first moment taken about the origin (x0 = 0)

$$m \equiv \bar{x} = \int_{-\infty}^{\infty} x f(x) dx$$

- The variance is the second moment taken about the mean

$$\sigma^2 = \overline{(x - \bar{x})^2} = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx$$

- The standard deviation is the square root of the variance

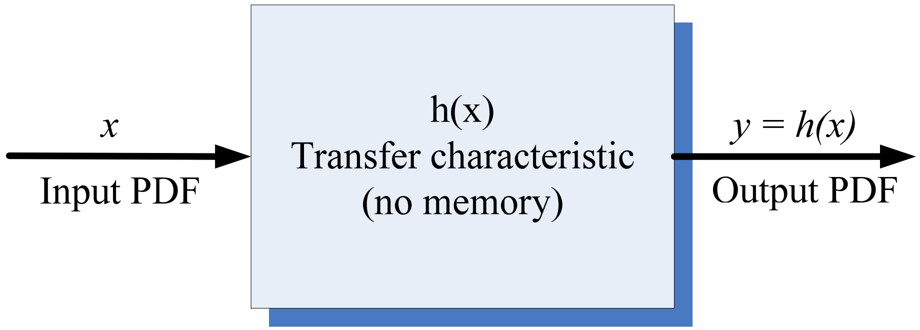
$$\sigma = \sigma^2 = \sqrt{\int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx}$$

- Theorem $\sigma^2 = \overline{x^2} - (\bar{x})^2$

# Some Useful Distributions

- Binomial Distribution
- Poisson Distribution
- Uniform Distribution
- Gaussian Distribution
- Sinusoidal Distribution

# Functional Transformations




Functional transformation of random variables

- Theorem If y =h(x), where h(.) is the output-to-input (transfer) characteristic of a device without memory, then the PDF of the output is

$$f_y(y) = \sum_{i=1}^{M} \frac{f_x(x)}{|dy/dx|}\Bigg|_{x=x_i=h_i^{-1}(y)}$$

- Where fx(x) is the PDF of the input x. M is the number of real roots of y = h(x). That is, the inverse of y = h(x) gives $x_1$, $x_2$, ..., $X_M$ for a single value of y.
- h(x) should not be confused with the impluse response of a linear network, which is denoted by h(t)

# Example

- Sinusoidal distribution

  Let y = h(x) = A sin x, Where x is uniformly distributed over -π to +π.

  First, -1 ≤ sin x ≤ 1, so f(y) = 0 for |y| > A.

  Second, there being two values of x, x1 and x2 for each value of y, since sin x = sin (π - x) applying functional transformation theorem,

$$f_y(y) = \begin{cases} \dfrac{f_x(x_1)}{|A\cos x_1|} + \dfrac{f_x(x_2)}{|-A\cos x_2|}, |y| \le A \\ 0, \qquad y \text{ elsewhere} \end{cases}$$

  By using the above result and the uniform PDF of fx(x),

$$f_y(y) = \begin{cases} \dfrac{1}{\pi\sqrt{A^2 - y^2}}, |y| \le A \\ 0, |y| > A \end{cases}$$

# Multi-Variable Statistics

- The N-dimensional CDF is $F(a_1, a_2, ..., a_N) = P[(x_1 \le a_1)(x_2 \le a_2)...(x_N \le a_N)]$

$$= \lim_{n \to \infty} \left[ \frac{n_{(x_1 \le a_1)(x_2 \le a_2)...(x_N \le a_N)}}{n} \right]$$

  Where the notation $(x_1 \le a_1)$ $(x_2 \le a_2)$… $(x_N \le a_N)$ is the intersection event consisting of the intersection of the events associated $x_1 \le a_1$, $x_2 \le a_2$, etc.

- The N-dimensional PDF is $f(x_1, x_2, ..., x_N) = \left. \frac{\partial^N F(a_1, a_2, ..., a_N)}{\partial a_1 \partial a_2 ... \partial a_N} \right|_{\mathbf{a}=\mathbf{x}}$

  Where $\mathbf{a}$ and $\mathbf{x}$ are the row vectors, $\mathbf{a} = (a_1, a_2, \ldots, a_N)$, $\mathbf{x} = (x_1, x_2, \ldots, x_N)$

- The expected value of y=h(**x**) is

$$\overline{[y]} = \overline{h(x_1, x_2, ..., x_N)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} h(x_1, x_2, ..., x_N) f(x_1, x_2, ..., x_N) dx_1 dx_2 ... dx_N$$

# Bi-variable Statistics

- The correlation (or joint mean) of x1 and x2 is

$$m_{12} = \overline{x_1 x_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2) dx_1 dx_2$$

- Two random variables x1 and x2 are said to be uncorrelated if

$$m_{12} = \overline{x_1 x_2} = \overline{x_1}\,\overline{x_2} = m_1 m_2$$

- Two random variables are said to orthogonal if $m_{12} = \overline{x_1 x_2} = 0$

- The covariance is

$$u_{12} = \overline{(x_1 - m_1)(x_2 - m_2)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - m_1)(x_2 - m_2) f(x_1, x_2) dx_1 dx_2$$

- The correlation coefficient is

$$\rho = \frac{u_{12}}{\sigma_1 \sigma_2} = \frac{\overline{(x_1 - m_1)(x_2 - m_2)}}{\sqrt{\overline{(x_1 - m_1)^2}}\sqrt{\overline{(x_2 - m_2)^2}}}$$

This is also called the normalized covariance. $-1 \le \rho \le +1$

# Multi-Variable Functional Transformation

Let $\mathbf{y} = \mathbf{h}(\mathbf{x})$ denote the transfer characteristic of a device (no memory) that has N inputs, denoted by $\mathbf{x} = (x_1, x_2, ..., x_N)$; N outputs, denoted by $\mathbf{y} = (y_1, y_2, ..., y_N)$; and $y_i = h_i(\mathbf{x})$. Furthermore, let $\mathbf{x}_i,\ i = 1, 2,\ ...,\ M$ denote the real roots (vector) of the equation $\mathbf{y} = \mathbf{h}(\mathbf{x})$.

The PDF of the outputs is then

$$f_{\mathbf{y}}(\mathbf{y}) = \sum_{i=1}^{M} \left. \frac{f_{\mathbf{x}}(\mathbf{x})}{|\,J(\mathbf{y}/\mathbf{x})\,|} \right|_{\mathbf{x}=\mathbf{x}_i=\mathbf{h}_i^{-1}(\mathbf{y})}$$

where $J(\mathbf{y}/\mathbf{x})$ is the Jacobian of the coordinat transformation to $\mathbf{y}$ from $\mathbf{x}$

$$J\left(\frac{\mathbf{y}}{\mathbf{x}}\right) = \text{Det}\begin{bmatrix} \frac{\partial h_1(\mathbf{x})}{\partial x_1} & \frac{\partial h_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial h_1(\mathbf{x})}{\partial x_N} \\ \frac{\partial h_2(\mathbf{x})}{\partial x_1} & \frac{\partial h_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial h_2(\mathbf{x})}{\partial x_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial h_N(\mathbf{x})}{\partial x_1} & \frac{\partial h_N(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial h_N(\mathbf{x})}{\partial x_N} \end{bmatrix}$$

where Det[.] denotes the determinant of the matrix

# Characteristic Function

- The characteristic function is defined as

$$M_x(v) \triangleq E[e^{jvx}] = \int_{-\infty}^{\infty} e^{jvx} f_x(x)dx$$

Mx(v) is the Fourier transform of fx(x), moment generating function

- For the nth moment,

$$E[x^n] = (-j)^n \left. \frac{\partial^n M_x(v)}{\partial v^n} \right|_{v=0}$$

- A pdf is obtained from the corresponding characteristic function by the inverse transform

$$f_x(x) = \int_{-\infty}^{\infty} e^{-jvx} M_x(v)dx$$

# The sum of two independent RVs

- Z = X + Y

  From the definition of the characteristic function of Z, we write

$$M_Z(v) \triangleq E[e^{jvZ}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{jv(x+y)} f_x(x) f_y(y)dxdy$$

$$= \int_{-\infty}^{\infty} e^{jvx} f_x(x)dx \int_{-\infty}^{\infty} e^{jvy} f_y(y)dy = E[e^{jvX}]E[e^{jvY}]$$

$$= M_X(v)M_Y(v)$$

The characteristic function is the Fourier transform of the corresponding PDF, and that a product in the frequency domain corresponds to convolution in the time domain, it follows that

$$f_z(z) = f_x(x) * f_y(y) = \int_{-\infty}^{\infty} f_x(z-u) * f_y(u)du$$

# Central Limit Theorem

- If we have the sum of a number of independent random variables with arbitrary one-dimensional PDFs, the central limit theorem states that the PDF for the sum of these independent random variable approaches a Guassian (Normal) distribution under very general conditions.
- Example

$y_1 = \sin(x_1)$

$y_2 = \sin(x_1)+\sin(x_2)$

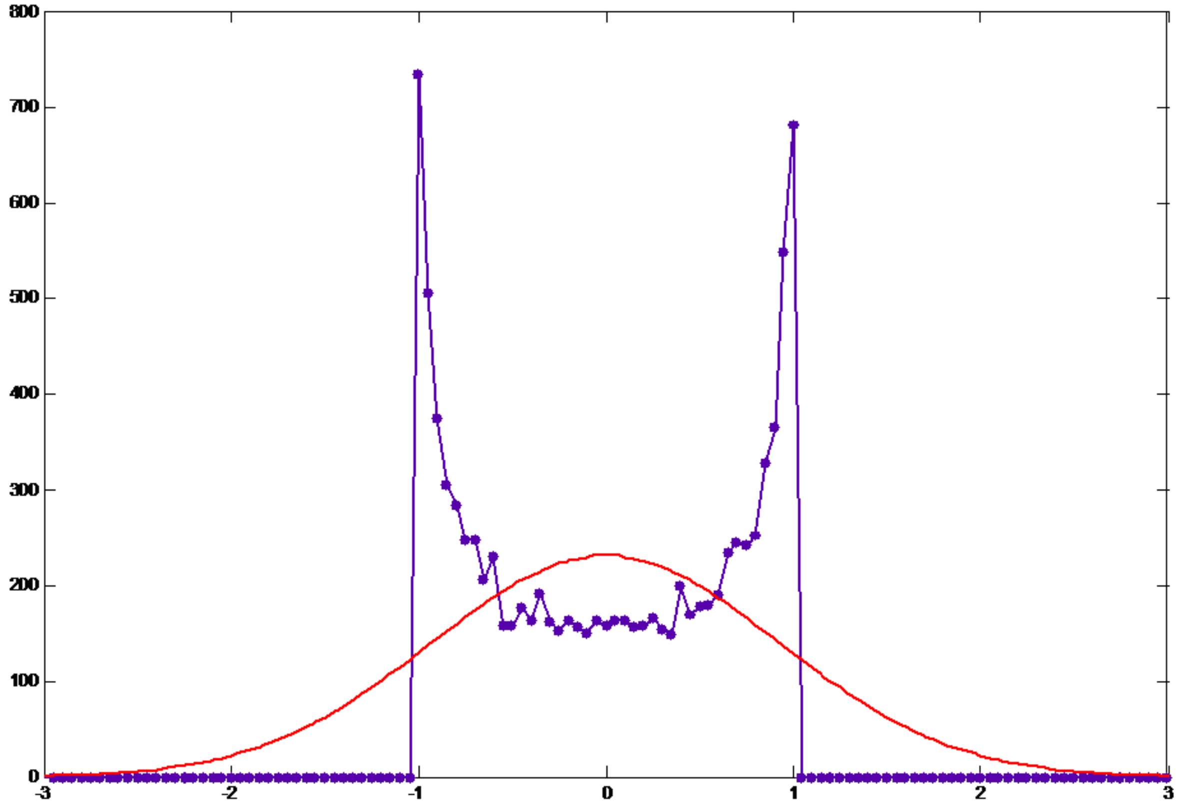
$y_4 = \sin(x_1)+\sin(x_2)+ \ldots + \sin(x_4)$

$y_8 = \sin(x_1)+\sin(x_2)+ \ldots + \sin(x_8)$

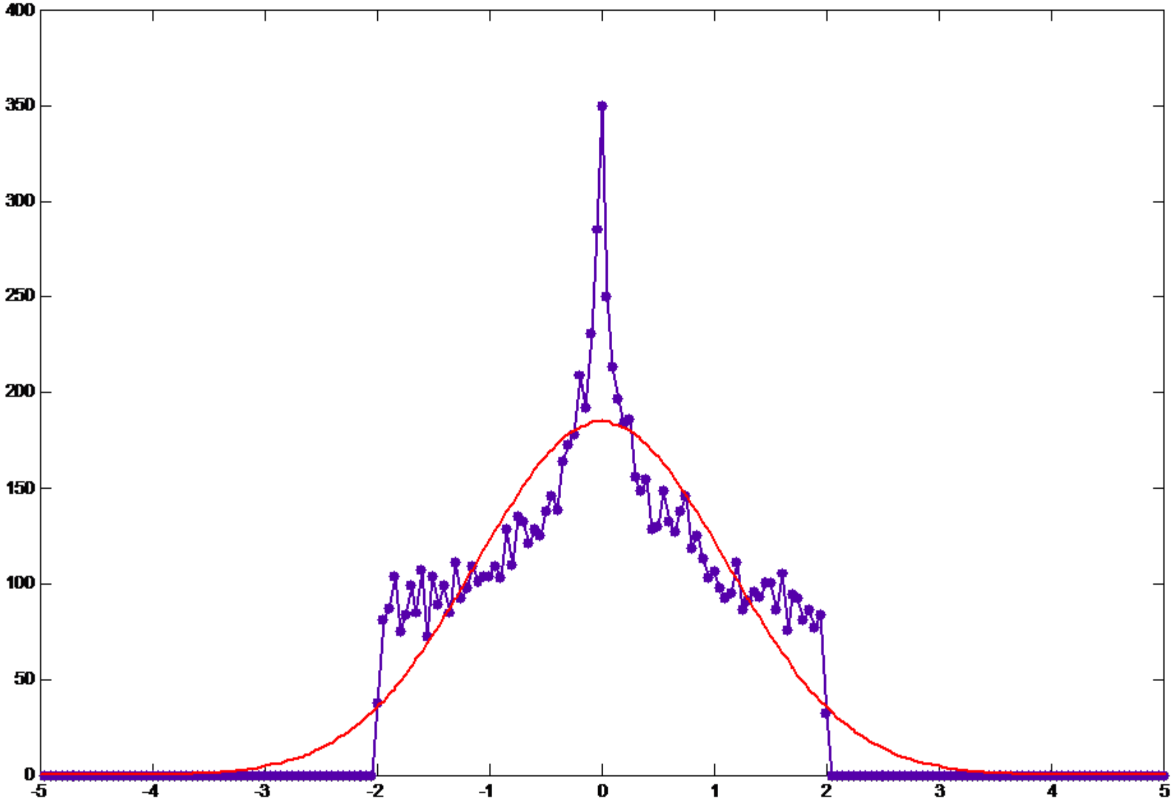
$y_{16} = \sin(x_1)+\sin(x_2)+ \ldots + \sin(x_{16})$

$x_1, x_2, \ldots, x_{16}$ are independent and uniformly distributed over 0 to $2\pi$.

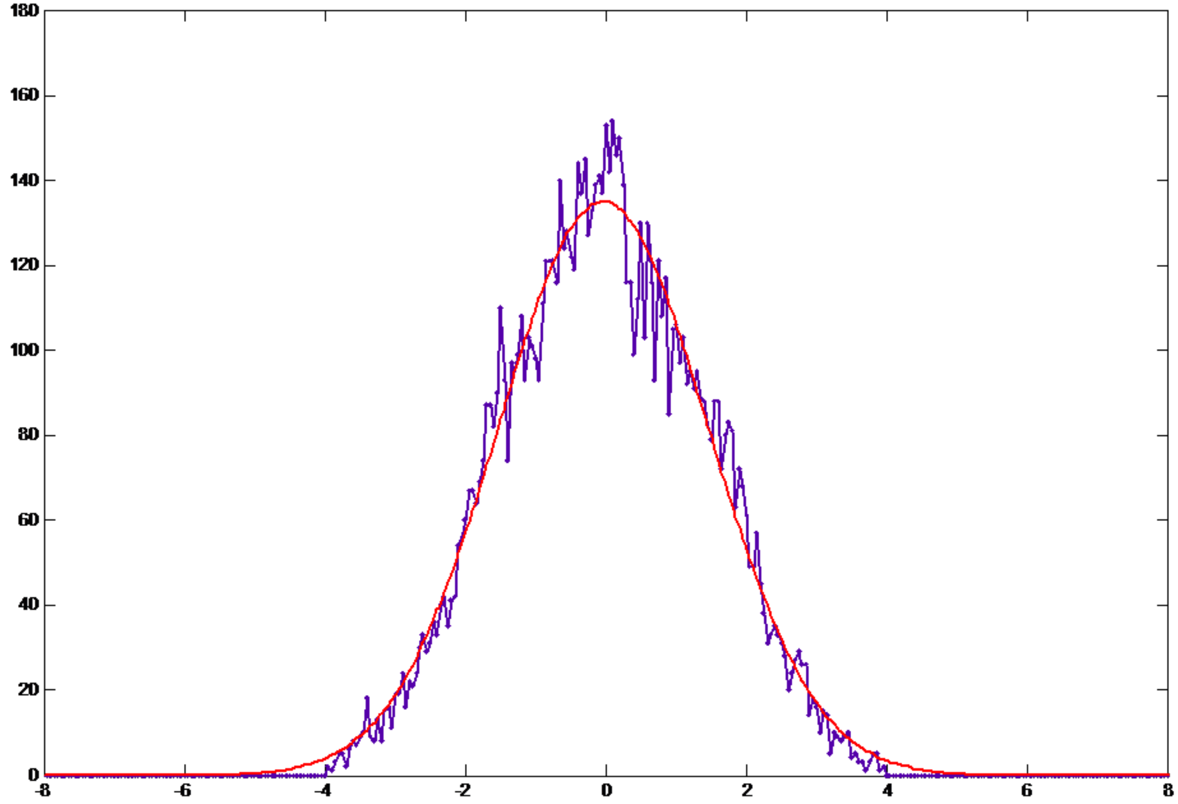
# N = 1



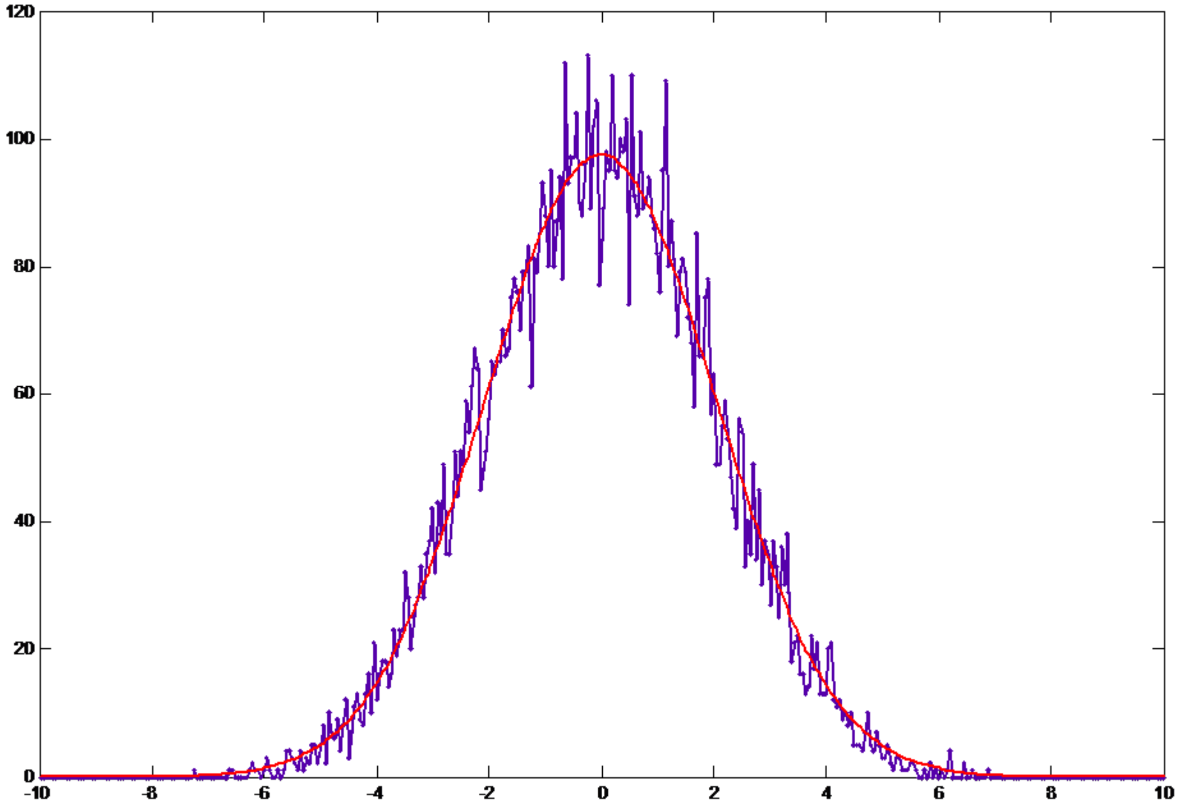
# N = 2



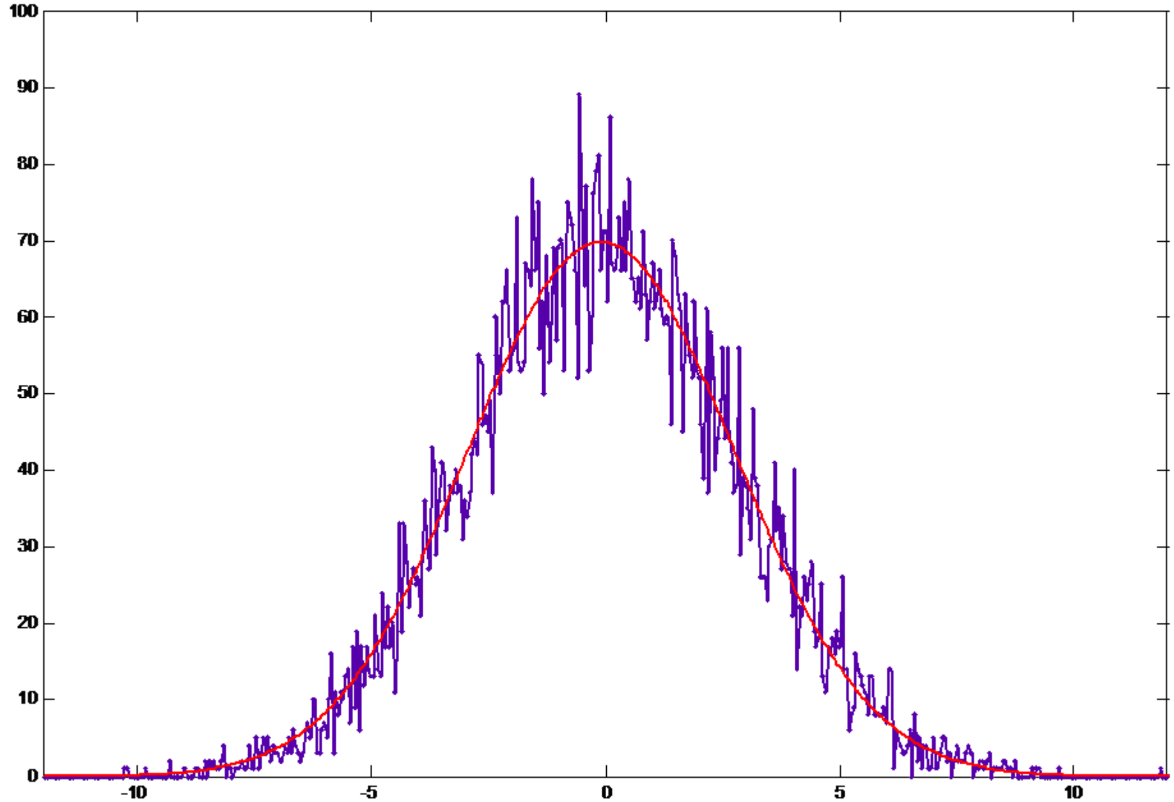
# N = 4



# N = 8



# N = 16



# Homework

- B-3, B-18, B-36, B-39, B-48