1 1-D STOCHASTIC FUNCTIONS – NOISE

• “noise” = function whose amplitude is derived from a random or a stochastic process (i.e., not deterministic)

• Deterministic: \( f \) at \( x \) specified completely by some parameters (width, amplitude, etc.)

• Stochastic function \( n \) at \( x \) selected from distribution that describes the probability of occurrence
  
  – Only statistical averages of signal amplitude are specified

• Example of “discrete” process: number of individual photons measured by detector in time interval \( \Delta t \)
  
  – must be integer
  
  – number counted during disjoint intervals of equal length exhibit a degree of “scatter” about some mean value.

• Example of “continuous” stochastic process
  
  – spatial distribution of photon arrivals from an optical system; the location \([x_n, y_n]\)

• Notation: \( n[x] \) instead of \( f[x] \); “\( n \)” is used to indicate its “noisy” character.
  
  – potential confusion because \( n \) also specifies an integer coordinate, e.g., \( f[n] \)

• Complex-valued noise constructed by adding two random variables at each coordinate after weighting one by \( i \):

\[
n(x) = n_1(x) + i n_2(x)
\]

1.1 PROBABILITY DISTRIBUTION

• \( P[n] \) = probability of the integer value \( n \) for a discrete distribution

• \( p[n] \) = probability density of a real-valued amplitude \( n \) derived from a continuous distribution

  – probability that continuous variable \( n \) lies within a specific interval is the integral of \( p[n] \) over that interval, e.g.

\[
P(a \leq n \leq b) = \int_a^b p[n] \, dn
\]

1.2 THE MOMENTS of a PROBABILITY DISTRIBUTION

• “distribution” of probability about origin of coordinates determined by moments of \( p[n] \)

• \( k^{th} \) moment is area of product of \( p[n] \) and a weight factor \( n^k \):

\[
m_k \{p[n]\} = \int_{-\infty}^{+\infty} n^k \, p[n] \, dn
\]
• “Zeroth moment” evaluated by setting \( k = 0 \); is projection of unit constant onto \( p[n] \), which must be the unit area of \( p[n] \):

\[
m_0 \{ p[n] \} = \int_{-\infty}^{+\infty} n^0 \ p[n] \ dn = \int_{-\infty}^{+\infty} p[n] \ dn = \langle n^0 \rangle = 1
\]

• First moment is projection of \( n^1 \) onto \( p[n] \)
  - “amplifies” contributions from large values of \( n \).
  - Mean value of amplitude \( n \) selected from probability distribution \( p[n] \)
  - Notations: \( \langle n \rangle \), \( \bar{n} \), and \( \mu \)

\[
m_1 \{ p[n] \} = \int_{-\infty}^{+\infty} n^1 \ p[n] \ dn = \langle n^1 \rangle = \langle n \rangle
\]

  - \( \langle n \rangle \) is value of \( n \) that divides \( p[n] \) into two equal parts

1.3 Central Moments

• expected values of \( (n - \langle n \rangle)^k \)

• measure weighted variation of \( p[n] \) about mean:

\[
\mu_k \{ p[n] \} \equiv \int_{-\infty}^{+\infty} (n - \langle n \rangle)^k \ p[n] \ dn
\]

• first central moment \( \mu_1 = 0 \).

• second moment is projection of \( n^2 \) onto \( p[n] \):

\[
m_2 \{ p[n] \} = \int_{-\infty}^{+\infty} n^2 \ p[n] \ dn \equiv \langle n^2 \rangle
\]

• second central moment of \( p[n] \) is variance \( \sigma^2 \)
  - measures “spread” of noise probability about mean value:

\[
\mu_2 \{ p[n] \} \equiv \int_{-\infty}^{+\infty} (n - \langle n \rangle)^2 \ p[n] \ dn \equiv \sigma^2
\]

• Expand \( (n - \langle n \rangle)^2 \) and evaluating three resulting integrals to obtain important relationship among the mean, variance, and \( \langle n^2 \rangle \)

\[
\sigma^2 = \langle n^2 \rangle - \langle n \rangle^2
\]
1.4 DISCRETE PROBABILITY LAWS

- model processes that have discrete (and often binary) outcomes
- particularly useful when constructing models of such imaging processes as photon absorption by sensor.
- Simplest type of discrete probability law applies to events that have only two possible outcomes
  - success or failure
  - true or false
  - on or off
  - head or tail.
- Individual implementation of binary event is “Bernoulli trial”
- Collective realizations of many such events described by binomial and Poisson distributions.

1.4.1 BERNOULLI TRIALS

- Consider flux of photons onto “imperfect” absorbing material
  - Individual photons may be absorbed or not
  - Testing of absorption of a photon is a Bernoulli trial
  - “successful” absorption indicated by 1, and “failure” by 0
  - Statistics of string of Bernoulli trials specified by probability of “success” (outcome “1”, denoted by “p”)
    \[ 0 \leq p \leq 1 \]
  - Probability of “failure” (outcome “0”) is \( 1 - p \), often denoted by “q”.
- Relative probabilities for particular absorber determined from physical model of interaction or from observed results of large number of Bernoulli trials
- “Images” of independent Bernoulli trials for different values of \( p \):

  Examples of 64 Bernoulli trials: (a) \( p = 0.2 \), (b) \( p = 0.5 \).
1.4.2 MULTIPLE BERNOULLI TRIALS – THE BINOMIAL PROBABILITY LAW

- $N$ Bernoulli trials where probability of outcome “1” is $p$ and the probability of outcome “0” is $1 - p$
- Number of possible distinct outcomes is $2^N$
- Relative likelihood of outcomes is determined by $p$.
  - Probability of specific sequence $101010 \cdots 10$ (alternating “1” and “0”, assuming $N$ is even)
    \[ p(10101010 \cdots 10) = p \cdot q \cdot p \cdot \cdots \cdot p \cdot q = p^{\frac{N}{2}} \cdot q^{\frac{N}{2}} \]
- Equal numbers of “successes” and “failures”.
- Probability of a different sequence where the first $\frac{N}{2}$ outcomes are “1” and the remaining $\frac{N}{2}$ outcomes are “0” is same:
  \[ p(111111 \cdots 000000) = p \cdot p \cdot p \cdot \cdots \cdot q \cdot q \cdot q = p^{\frac{N}{2}} \cdot q^{\frac{N}{2}} \]
- Two distinct outcomes have same number of “successes” and “failures”, and therefore have identical histograms.
- Outcome of $N$ trials with exactly $n$ “successes”. The probability of a specific such outcome is:
  \[ p_1 \cdot p_2 \cdot p_3 \cdots \cdot p_n \cdot q_1 \cdot q_2 \cdot q_3 \cdots \cdot q_{N-n} = p^n \cdot q^{N-n} \]
- In many applications, order of arrangement is not significant; only total number of successes $n$ matters.
- Compute number of possible combinations of $n$ successes in $N$ trials.
- It is straightforward to show that this number is the “binomial coefficient”:
  \[ \frac{N!}{(N-n)! \cdot n!} = \binom{N}{n} \]
  - probability of $n$ “successes” in $N$ trials is:
    \[ P_n = \frac{N!}{n! \cdot (N-n)!} \cdot p^n \cdot (1-p)^{N-n} = \binom{N}{n} \cdot p^n \cdot [1 - p]^{N-n} \]
- Consider a coin flip with two outcomes, $H$ and $T$.
  - $N = 4 \implies$ number combinations with $n = 2$ heads is $\frac{4!}{2! \cdot 2!} = 6$
    - $HHTT, HTHT, THHT, THTH, HTTH, \text{ and } TTHH$
  - If $H$ and $T$ equally likely ($p = q = 0.5$) \implies probability of two heads in four flips is $P_2 = 6(0.5)^2(0.5)^2 = 0.375$.
  - Number of realizations of no heads in four flips is $\frac{4!}{0! \cdot 4!} = 1$, Probability of no heads in four flips is $P_0 = 1(0.5)^0(0.5)^4 = 0.0625$.
  - Binomial law also applies to cases where $p \neq q$
    - Flipping “unfair” coin
    - $p = 0.75 \implies$ probability that four flips would produce two heads is $P_2 = 6(0.75)^2(0.25)^2 \approx 0.211$
• less than probability of two heads in four flips of fair coin.

• Mean and variance of number of outcomes with individual probability \( p \) in experiments obtained by substitution

\[
\langle n \rangle = N p \\
\sigma^2 = N p \ [1 - p]
\]

• “Shape” of histogram is approximately Gaussian.

  – Gaussian distribution is continuous
  – Binomial distribution is discrete
  – Observation suggests that samples obtained from large number of independent Bernoulli trials with probability \( p \) may be approximately generated by thresholding values generated by a Gaussian distribution.
Realizations of $N$ Bernoulli trials at 8192 samples with $p = 0.5$ and the resulting histograms: (a) $N = 1$ trial per sample, two (nearly) equally likely outcomes; (b) $N = 2$; (c) $N = 10$; and (d) $N = 100$. The histogram approaches a Gaussian function for large $N$. 
1.4.3 POISSON PROBABILITY LAW

- Approximation to binomial law for large numbers of rarely occurring events, i.e., $N >> 1$ and $p \to 0$
- Mean number of events is $\langle n \rangle = Np$, denoted by $\lambda$
- Form of Poisson law obtained by substituting into the binomial law in limit $N \to \infty$:

$$p_n = \lim_{N \to \infty} \left\{ \left( \begin{array}{c} N \\ n \end{array} \right) \left[ \frac{\lambda}{N} \right]^n \left[ 1 - \frac{\lambda}{N} \right] [N-n] \right\}$$

- Take the natural logarithm of both sides to obtain:

$$\log_e [p_n] = \lim_{N \to \infty} \left\{ \log_e \left[ \left( \begin{array}{c} N \\ n \end{array} \right) \left[ \frac{\lambda}{N} \right]^n \left[ 1 - \frac{\lambda}{N} \right] [N-n] \right] \right\}$$

$$= \lim_{N \to \infty} \left\{ \log_e \left[ \left( \begin{array}{c} N \\ n \end{array} \right) \left[ \frac{\lambda}{N} \right]^n \right] \right\} + \lim_{N \to \infty} \left\{ [N-n] \log_e \left[ 1 - \frac{\lambda}{N} \right] \right\}$$

- Use fact that $n$ is small to evaluate first additive term:

$$\lim_{N \to \infty} \left\{ \log_e \left[ \frac{N!}{(N-n)!n!} \left[ \frac{\lambda}{N} \right]^n \right] \right\} \simeq \log_e \left\{ \frac{N^n}{n!} \left( \frac{\lambda}{N} \right)^n \right\}$$

- Second term evaluated by recognizing as ratio of two terms that both approach zero in the limit

- Apply l’Hôpital’s rule:

$$\lim_{N \to \infty} \left\{ \frac{\log_e (1 - \frac{\lambda}{N})}{(N-n)^{-1}} \right\} = \lim_{N \to \infty} \left\{ \frac{\frac{d}{dN} \log_e (1 - \frac{\lambda}{N})}{\frac{d}{dN} (N-n)^{-1}} \right\}$$

$$= \lim_{N \to \infty} \left\{ \frac{(1 - \frac{\lambda}{N})^{-1} \left( \frac{\lambda}{N} \right)}{-(N-n)^{-2}} \right\}$$

$$= \lim_{N \to \infty} \left\{ -\lambda \left( \frac{N-n}{N} \right)^2 \left( 1 - \frac{\lambda}{N} \right)^{-1} \right\}$$

- Collect terms:

$$\log_e [p_n] = \log_e \left[ \frac{\lambda^n}{n!} \right] - \lambda \implies p_n = \left[ \frac{\lambda^n}{n!} \right] e^{-\lambda}$$

- Poisson distribution is particular limiting case of binomial distribution

- Mean, variance, and third central moment of Poisson distribution are identically $\lambda$
Comparison of binomial and Poisson random variables, $N = 100$: (a) binomial, $p = 0.75$, $\langle n \rangle = 75.05$, $\sigma^2 = 18.68$; (b) Poisson, $\lambda = 75$, $\langle n \rangle = 74.86$, $\sigma^2 = 74.05$; (c) binomial, $p = 0.25$, $\langle n \rangle = 24.93$, $\sigma^2 = 18.77$; (d) Poisson, $\lambda = 25$, $\langle n \rangle = 25.01$, $\sigma^2 = 24.85$; (e) binomial, $p = 0.05$, $\langle n \rangle = 5.00$, $\sigma^2 = 4.71$; (f) Poisson, $\lambda = 5$, $\langle n \rangle = 4.97$, $\sigma^2 = 4.97$. 
1.5 CONTINUOUS PROBABILITY DISTRIBUTIONS

1.5.1 UNIFORM DISTRIBUTION

- generates most intuitive type of noise
  - amplitude \( n \) equally likely to occur within any finite interval of equal size.
    
    \[
    p_{\text{Uniform}}[n] = \frac{1}{|b|} \cdot \text{RECT} \left( \frac{n - \langle n \rangle}{b} \right)
    \]
    
    - \( b \) is width of allowed values of \( n \)
    - \( \langle n \rangle \) is mean value
    - Multiplicative scale factor \( b^{-1} \) ensures unit area
    - Variance of uniform distribution is \( \sigma^2 = \frac{b^2}{12} \).

Uniformly distributed random variable on interval \([0,1)\) with \( \mu = 0.5 \) and \( \sigma^2 = \frac{1}{12} \): (a) sample, (b) histogram.
1.5.2 NORMAL DISTRIBUTION

- Familiar symmetric “bell curve” of probability
- Most applicable of all probability laws
- $\langle n \rangle = \text{most likely amplitude (peak of the probability density)}$
- Probability that amplitude will differ from mean progressively decreases as the value moves away from $\langle n \rangle$
- Probability density function is Gaussian function with width parameter $b$ proportional to standard deviation $\sigma$ of probability distribution

$$p_{\text{Normal}}[n] = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(n-\langle n \rangle)^2}{2\sigma^2}}$$

- Leading factor $(\sqrt{2\pi})^{-1}$ ensures that area of the probability density function is unity

Samples of a random variable generated by a normal distribution with $\langle n \rangle = 0$, $\sigma^2 = 1$: (a) samples, (b) histogram.

Central-Limit Theorem:

- Cascade of stochastic processes derived from (nearly) arbitrary set of probability density functions generates a normal distribution
- Central-limit theorem ensures that probability law of outputs is generally Gaussian, to good approximation.
1.5.3 RAYLEIGH DISTRIBUTION

- Imaging applications that involve Fourier transforms of distributions of complex-valued random variables
  - description of Fraunhofer diffraction from a random scatterer
  - computer-generated holography.

- Distribution of magnitude where real and imaginary parts are random variables selected from same Gaussian distribution.

- Probability density function characterized by single parameter $a$:
  
  \[ P_{\text{Rayleigh}}[n] = \frac{n}{a^2} e^{-\left(\frac{n^2}{2a^2}\right)} \text{STEP}[n] \]

- STEP function ensures that allowed amplitudes $n$ must be nonnegative

- Mean $\langle n \rangle$ and variance $\sigma^2$ of the Rayleigh distribution must be functions of the parameter $a$:
  
  \[ \langle n \rangle = \sqrt{\frac{\pi}{2}} a \approx 1.25a \]
  \[ \sigma^2 = \left(2 - \frac{\pi}{2}\right) a^2 \approx 0.429a^2, \quad \sigma \approx 0.655a \]

Rayleigh-distributed random variable generated from Gaussian-distributed random variables in quadrature, each with $\langle n \rangle = 0, \sigma^2 = 1$: (a) sample, (b) histogram. Resulting mean and variance are $\langle n \rangle = 1.270, \sigma^2 = 0.424$. 