ECE 3075A

Random Signals

Lecture 21

Review of Random Variables and Statistics; Concept of Random Processes

School of Electrical and Computer Engineering
Georgia Institute of Technology
Fall, 2003

Expectation

- Let $\{p_i\}_{i=1}^{L}$ be the (a priori) probability associated with a source which puts out symbols $\{X = x_i\}_{i=1}^{L}$;
- Define information in symbol x_i as $-\log_2 p_i$;
- The average information is called entropy and is defined as

$$H = E_X\left[-\log_2 p_i\right] = -\sum_{i=1}^{L} p_i \log_2 p_i$$

Example:
A source sends out signal according to the outcome of a 2-coin throwing experiment. The signal symbol set consists of HH, {HT or TH}, and TT – that is, no distinction between HT and TH is made. The two coins are biased; one has a head probability of 0.3 and the other 0.6. Determine the entropy of this signal source.

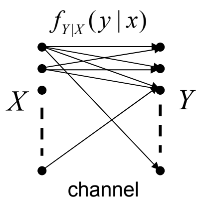
$$\Pr\{\text{HH}\} = 0.3\times 0.6 = 0.18, \quad \Pr\{\text{HT or TH}\} = \Pr\{\text{HT}\} + \Pr\{\text{TH}\} = 0.3\times 0.4 + 0.7\times 0.6 = 0.54, \quad \Pr\{\text{TT}\} = 0.7\times 0.4 = 0.28$$

$$H = -\sum_{i=1}^{L} p_i \log_2 p_i = -0.18\log_2 0.18 - 0.54\log_2 0.54 - 0.28\log_2 0.28 = 1.4396 \text{ bits/symbol}$$

If HT and TH were considered different signals, the entropy would have been 1.8522 bits/symbol, a difference of 0.4126 bits/symbol.

Conditional Density Functions

- Consider a "random" source which puts out signal X, which at the destination is observed as Y due noise and distrotion.
- We are often interested in estimating x (i.e., what was sent) based on the received signal y.



Example: Use the previous 2-coin source.

$$\Pr\{X = \text{HH}\} = 0.3\times 0.6 = 0.18, \quad \Pr\{X = \text{HT or TH}\} = 0.54, \quad \Pr\{X = \text{TT}\} = 0.28$$

Suppose the presence of channel noise causes a confusion between H and T, with probability:

$$\Pr\{\text{H}\,|\,\text{T}\} = 0.1, \quad \Pr\{\text{T}\,|\,\text{H}\} = 0.05$$

x	y	$f_{Y\|X}(y\|x)$	$f_{Y,X}(y,x)$	$f_{X\|Y}(x\|y)$
HH	HH	0.9025		
HH	HT, TH	0.095	0.0171	0.0321
HH	TT	0.0025	0.00045	
HT,TH	HT,TH			0.8731
HT,TH	HH			
HT,TH	TT	0.045	0.0243	
TT	TT	0.81	0.2268	0.9016
TT	TH,HT			0.0948
TT	HH		0.0028	

$$\Pr\{y = \text{HH}\} = \sum_x f_{Y|X}(y\,|\,x)\Pr\{x\} = 0.21655$$

$$\Pr\{y = \text{HT or TH}\} = ??$$

$$\Pr\{y = \text{TT}\} = ??$$

The entropy at the destination is
$H = 1.4633$ bits/symbol

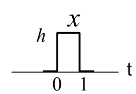
Sum of Random Variables

Let $V_i,\ i = 1, 2, \cdots, L$ be random variables with zero mean and variances $\sigma_{V_1}^2, \sigma_{V_2}^2, \cdots, \sigma_{V_L}^2$, respectively.

a. Show that their sum $V = V_1 + V_2 + \cdots + V_L$ has a variance $\sigma_V^2 = \sigma_{V_1}^2 + \sigma_{V_2}^2 + \cdots + \sigma_{V_L}^2$ that is, the variance of sum is sum of variances.

Let x be an ideal impulse signal, as shown, which has power h^2. In reality, however, the level of this pulse signal often fluctuates due to noise, such as any of the V_i above. Assume $\sigma_{V_1}^2 = \sigma_{V_1}^2 = \cdots = \sigma_{V_L}^2 = \sigma^2$. We say h^2/σ^2 is the signal to noise ratio (SNR or S/N) of the signal, in that the variance of the zero-mean random noise is used as the noise power. To fight against the noise, in signal transmission, we often repeatedly sent the pulse (with noise when it reaches the destination) L times so that the receiver can have a better reception by adding the received (noisy) pulses, $Y_i = x + V_i$, together to form $Y = Y_1 + Y_2 + \cdots + Y_L$.

b. Determine the SNR in Y.
c. How many repetitions do we need to transmit in order to have a gain of 3 dB in SNR at the receiver?



Functions of Random Variables

X and Y are independent exponential random variables with common parameter λ. Define $U = X + Y$ and $V = X - Y$.

1. Find $f_{U,V}(u,v)$
2. Find the marginal density functions $f_U(u)$ and $f_V(v)$
3. Find the correlation between U and V.

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\lambda^2} e^{-(x+y)/\lambda}, \quad x > 0,\ y > 0$$

Since $|v| < u$ (due to the fact that x and y are always positive), only one solution for x and y needs to be considered.

$$J = \begin{vmatrix} \frac{\partial U}{\partial X} & \frac{\partial U}{\partial Y} \\ \frac{\partial V}{\partial X} & \frac{\partial V}{\partial Y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \qquad \therefore\ f_{U,V}(u,v) = \frac{1}{2\lambda^2} e^{-u/\lambda}, \quad 0 < |v| < u < \infty$$

$$f_U(u) = \int_{-u}^{u} f_{U,V}(u,v)\, dv = \frac{u}{\lambda^2} e^{-u/\lambda}, \quad 0 < u < \infty$$

$$f_V(v) = \int_{|v|}^{\infty} f_{U,V}(u,v)\, du = \frac{1}{2\lambda} e^{-|v|/\lambda}, \quad -\infty < v < \infty$$

$$E[UV] = E[(X+Y)(X-Y)] = E[X^2] - E[Y^2] = 0$$

Curve Fitting

- Observations are made on two or more variables that may have a certain relationship – as in study of physics, for example. These data points, when plotted in a multi-dimensional space, form a scatter diagram.
- We use curve fitting to find a mathematical relationship, often simplified and parameterized, among the variables. The mathematical equation that relates the variables are called the regression equation which defines a regression curve, which fits the given observation data in some optimal sense, so-called criterion of goodness-of-fit.

Given n data points, in a 2-D example, $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ we choose to fit a curve define by $y = g(x)$ to the data such that the fitting error defined as $D = \sum_{i=1}^{n} [y_i - g(x_i)]^2$ is minimized. This leads to a **least-square regression** curve. We can choose other criterion if we like.

Least-Square Regression

We perform linear regression on a data set $\{(x_i, y_i)\}_{i=1}^{n}$ where x_i represents the input to a system $g(\bullet)$ and y_i is the output, observed with noise, $y_i = g(x_i) + v_i$. With linear regression, $y = g(x) = a + bx$.

Let $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$

The least square result is $b = \dfrac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - (\overline{x})^2}$ $\quad a = \overline{y} - b\overline{x}$

Example:

x	1.2	1.5	0.4	0.7	3.2	-0.4	0.1	1.1	-0.3	1.0	0.9	2.1
y	3.0	3.7	1.1	1.3	6.5	-1.2	-0.1	2.9	-0.9	2.0	2.0	4.5
$G(x)$	2.598	3.257	0.84	1.499	6.991	-.917	0.181	2.378	-.698	2.158	1.939	4.575
v'	0.402	0.443	0.260	-.199	-.491	-.283	-.281	0.522	-.202	-.158	0.061	-.075

$$y = g(x) = a + bx \qquad v'_i = y_i - g(x_i)$$

$$b = 2.197, \quad a = -0.0387 \qquad \overline{v'} \approx 0, \quad \sigma_{v'} = 0.318$$

Random Process and Outcomes of Experiment

- As in the random variable case which results from associating a random event with a set on the real line, a random process can be viewed as such a concept enlarged to include ***time***.
- We thus assign a time function to every outcome ξ
  $x(t, \xi)$
- The family of all such functions $X(t, \xi)$, with the corresponding outcomes forming a probability space, is called a random process. A short form notation is

$$x(t,\xi) \Rightarrow x(t)$$
$$X(t,\xi) \Rightarrow X(t)$$

Random Process

- A random process is a time function, the value of which at any instance of time is a random variable.
- Concepts and Notation:
  - Random Process: $X(t)$, an ensemble of functions
  - Sample function: $x(t)$

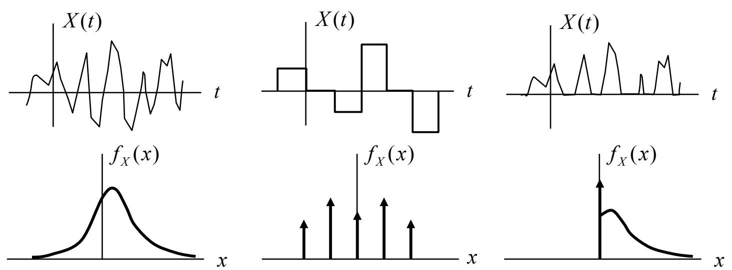
    A sample function is a realization of a random process, which implies an ensemble of functions that collectively form the observation space. A sample function is an element in the observation space.
  - At a specific time t_1, $X(t_1) = X_1$ is a random variable.

Types of Processes

- Discrete ⇔ Continuous
- Deterministic ⇔ Non-deterministic
- Stationary ⇔ Non-stationary
- Ergodic ⇔ Non-ergodic

Continuous and Discrete Random Process

- If $X(t_i) = X_i$ for any t_i is a continuous random variable, $X(t)$ is a continuous random process.
- If $X(t_i) = X_i$ for any t_i is a discrete random variable, $X(t)$ is a discrete random process.
- If $X(t_i) = X_i$ has a mixed (continuous and discrete) distribution, $X(t)$ is a mixed random process.



Continuous Discrete Mixed