ECE 3075A
Random Signals

Lecture 6
Multiple Random Variables, Joint Distributions, and Correlation

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

Some unfinished topics around single random variables

Entropy – Average Information

- 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

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

For example, if $L = 2$, and $x_1 = 0$, and $x_2 = 1$, then it is a binary source with symbols 0 and 1. And if $p_1 = p_2 = 0.5$, then every symbol the source puts out carries one **bit** of information.
if $p_1 \neq p_2$, say, $p_1 = 0.8$ and $p_2 = 0.2$, then the average information in each symbol is less than one bit.

Conditional Probability Distribution

We define the conditional probability the same as before.

$$F_X(x \mid M) = \Pr(X \le x \mid M) = \frac{\Pr(X \le x, M)}{\Pr(M)}, \quad \Pr(M) > 0$$

If we use the event mapping concept, $\Pr(X \le x, M)$ is the probability of all the outcomes which realize both events $X(\xi) \le x$ and $\xi \in M$.

If $M = \{X \le m\}$, $F_X(x \mid M) = \Pr(X \le x \mid X \le m) = \dfrac{\Pr(X \le x, X \le m)}{\Pr(X \le m)}$

If $x \le m$, $F_X(x \mid M) = \dfrac{\Pr(X \le x)}{\Pr(X \le m)} = \dfrac{F_X(x)}{F_X(m)}$ If $x \ge m$, $F_X(x \mid M) = \dfrac{\Pr(X \le m)}{\Pr(X \le m)} = 1$

Conditional probability density function has all the properties of a usual pdf. $f_X(x \mid M) = \dfrac{dF_X(x \mid M)}{dx}$

Conditional Expected Value

- The expected value of X, given event M, is

$$E[X \mid M] = \int_{-\infty}^{\infty} x f_X(x \mid M)\,dx$$

if $M = \{X \le m\}$, $f_X(x \mid X \le m) = \begin{cases} \dfrac{f_X(x)}{F_X(m)}, & x < m \\ 0, & x \ge m \end{cases}.$

Thus, $E[X \mid X \le m] = \dfrac{\int_{-\infty}^{m} x f_X(x)\,dx}{\int_{-\infty}^{m} f_X(x)\,dx}$

which is the expected value of X when X is constrained to the set/event $\{X \le m\}$.

Useful & Intriguing Inequalities

- Chebychev Inequality

$$\Pr\{|X - \bar{X}| \ge \varepsilon\} \le \sigma_X^2 / \varepsilon^2$$

$$\Pr\{|X - \bar{X}| \ge \varepsilon\} = \int_{-\infty}^{\bar{X}-\varepsilon} f_X(x)\,dx + \int_{\bar{X}-\varepsilon}^{\infty} f_X(x)\,dx = \int_{x \in B} f_X(x)dx$$

where $B = \{R - \{|x - \bar{X}| < \varepsilon\}\}$

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x-\bar{X})^2 f_X(x)\,dx \ge \int_{x\in B} (x-\bar{X})^2 f_X(x)\,dx \ge \varepsilon^2 \int_{x \in B} f_X(x)\,dx$$
$$= \varepsilon^2 \Pr\{|X - \bar{X}| \ge \varepsilon\}$$

- Markov Inequality: for non-negative r.v. X,

$$\Pr\{X \ge a\} \le E[X]/a, \quad a > 0$$

Monte Carlo Simulation

- To generate a random variable with prescribed probability distribution function, often needed in computer simulation of a certain physical phenomenon such as noise in a measurement;
- Construction of such a simulation begins with a random variable that has uniform distribution on [0,1];
- Find a monotonically non-decreasing transformation $Y=T(X)$ such that Y has the prescribed distribution $F_Y(y)$.

$$\Pr\{Y \le y\} = F_Y[y = T(x)] = F_X(x) = \Pr\{X \le x\}$$

Since X is uniformly distributed, $F_X(x) = x$ for $0 \le x \le 1$.
Transformation $y = T(x) = F_Y^{-1}(x)$, for $0 \le x \le 1$ will produce the desired simulation of r.v. Y.

Uniform Random Number Generator

- Algorithmic generation of *random* numbers that display uniform distribution between 0 and 1;
- No algorithm can produce true random numbers; they are called pseudo-random numbers instead;
- Integer linear congruence method is often used:

$$x_{n+1} = ((ax_n + c))_m = (ax_n + c)_{\bmod m}$$

where a and c are certain fixed values and x_0 is an initial number known as the seed. Normalizing the result by m will produce a pseudo-random number sequence with value in [0,1). Such a sequence is periodic with a “cycle” less than or equal to m.

Example of Monte Carlo Simulation

Let $a = 9, c = 5, m = 16,$ and $x_0 = 4.$

$$x_{n+1} = (9x_n + 5)_{\text{mod } 16}$$

Try $a = 9, c = 3, m = 16,$ and $x_0 = 4,$ and see what happens.

Integer sequence: 4, 9, 6, 11, 8, 13, 10, 15, 12, 1, 14, 3, 0, 5, 2, 7, 4, ….

Normalized plus an epsilon (0.0001):
0.2501, 0.5626, 0.3751, 0.6876, 0.5001, 0.8126, 0.6251, 0.9376, 0.7501, 0.0626, 0.8751, 0.1876, 0, 0.3126, 0.1251, 0.4376, 0.2501, …

E.g., Desired Probability Distribution Function: $F_Y(y) = \dfrac{1}{1+e^{-ay}}$

$$F_Y(y) = \frac{1}{1+e^{-ay}} = x, \quad y = -\frac{1}{a}\ln\left(\frac{1}{x} - 1\right)$$

y = -0.54904, 0.12586, -0.2552, 0.394461, 0.0002, 0.733497, 0.255626, 1.354879, 0.549573, -1.35317, 0.973412, -0.73284, -4.60512, -0.394, -0.9725, -0.12545, -0.54904, …..

What is the corresponding probability density function $f_Y(y)$?

Another Example of RNG

$$s = 2111111111x_{n-4} + 1492x_{n-3} + 1776x_{n-2} + 5115x_{n-1} + c$$

a 64-bit unsigned integer

$$x_n = (s)_{\text{mod } 2^{32}}$$

where $c = \left\lfloor \dfrac{s}{2^{32}} \right\rfloor$ = largest integer less than or equal to $\dfrac{s}{2^{32}}$

The s in calculating the floor above is the previous value (during calculation of x_{n-1}).

Specifications:

- 32-bit integer output
- cycle length is $3 \cdot 10^{47}$

Characteristic Function

- Definition

$$\Phi_X(u) = E[e^{juX}] = \int_{-\infty}^{\infty} e^{jux} f_X(x)\,dx$$

that is, the characteristic function of a random variable can be viewed as the Fourier transform of its probability density function.

- The pdf is then the inverse Fourier transform of the characteristic function

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-jux}\Phi_X(u)\,du$$

$$\frac{d}{du}\Phi(u) = \int_{-\infty}^{\infty}\left(\frac{d}{du}e^{jux}\right) f_X(x)\,dx = \int_{-\infty}^{\infty} jxe^{jux} f_X(x)\,dx$$

$$\left.\frac{d}{du}\Phi(u)\right|_{u=0} = \int_{-\infty}^{\infty} jx\, f_X(x)\,dx = j\overline{X} \qquad \left.\frac{d^n}{du^n}\Phi(u)\right|_{u=0} = j^n E[X^n] = j^n\overline{X^n}$$

Main Subjects of This Lecture

Two Random Variables

- As in combined experiments, we are interested in the co-occurrence of separate events.

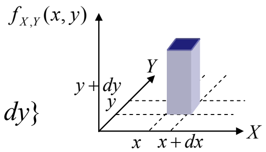
  Try to recall and answer: Do they have to be simultaneous? What is the difference, if any, between tossing a die and a coin together and tossing them separately? What is a combined event in the Cartesian product observation/sample space? What is the definition of **independent** experiments? And, likewise, **dependent** experiment?

- **Joint Probability Distribution Function** of two r.v., X and Y $\quad F_{X,Y}(x,y)=\Pr\{X\le x, Y\le y\}$

  1. $0\le F_{X,Y}(x,y)\le 1, \quad -\infty<x<\infty, -\infty<y<\infty$
  2. $F_{X,Y}(-\infty,y)=F_{X,Y}(x,-\infty)=F_{X,Y}(-\infty,-\infty)=0$
  3. $F_{X,Y}(\infty,\infty)=1$
  4. $F_{X,Y}(x,y)$ is a non-decreasing function of either x or y
  5. $F_{X,Y}(\infty,y)=F_Y(y), \quad F_{X,Y}(x,\infty)=F_X(x)$

Joint Probability Density Functions

Definition $\quad f_{X,Y}(x,y)=\dfrac{\partial^2 F_{X,Y}(x,y)}{\partial x\,\partial y}$



$$f_{X,Y}(x,y)\,dx\,dy=\Pr\{x<X\le x+dx, y<Y\le y+dy\}$$

1. $f_{X,Y}(x,y)\ge 0, \quad -\infty<x<\infty, -\infty<y<\infty$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy=1$
3. $F_{X,Y}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du$
4. $f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy, \quad f_Y(y)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \quad \rightarrow$ marginals
5. $\Pr\{x_1<X\le x_2, y_1<Y\le y_2\}=\int_{x_1}^{x_2}\int_{y_1}^{y_2} f_{X,Y}(x,y)\,dy\,dx$

Example

The joint density function of X and Y is given by

$$f_{X,Y}(x,y)=u(x)u(y)xe^{-x(y+1)}$$

Find the marginal density functions.

$$f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy=\int_0^{\infty} u(x)xe^{-x(y+1)}dy$$

$$=u(x)xe^{-x}\int_0^{\infty} e^{-xy}dy=u(x)xe^{-x}\left.\frac{e^{-xy}}{-x}\right|_{y=0}^{\infty}=u(x)e^{-x}$$

$$f_Y(y)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx=\int_0^{\infty} u(y)xe^{-x(y+1)}dx$$

$$=u(y)e^{-x(y+1)}\left[\frac{x}{-(y+1)}-\frac{1}{(y+1)^2}\right]_{x=0}^{\infty}=\frac{u(y)}{(y+1)^2}$$

Expectation of Functions of Two R.V.s

- Similar definition as in single random variable case

$$E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)\,dx\,dy$$

- When $\quad g(X,Y)=X^nY^k$

  $E[X^nY^k]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^ny^k f_{X,Y}(x,y)\,dx\,dy$ is called the ***joint moment of*** X ***and*** Y. When $n=k=1$, it is called correlation. $n+k$ is the order of the moment.

$$E[(X-\overline{X})(Y-\overline{Y})]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\overline{X})(y-\overline{Y})f_{X,Y}(x,y)\,dx\,dy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy-\overline{X}y-x\overline{Y}+\overline{X}\overline{Y})f_{X,Y}(x,y)\,dx\,dy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\,dx\,dy-\overline{X}\overline{Y}=\overline{XY}-\overline{X}\overline{Y}$$

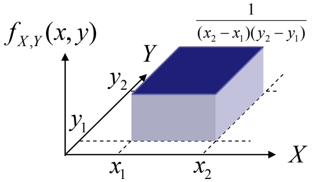
Two Uniform Random Variables

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{(x_2 - x_1)(y_2 - y_1)}, & x_1 < x \le x_2,\ y_1 < y \le y_2 \\ 0, & \text{elsewhere} \end{cases}$$

Marginals

$$f_X(x) = \int_{y_1}^{y_2} \frac{1}{(x_2-x_1)(y_2-y_1)}\,dy = \frac{y}{(x_2-x_1)(y_2-y_1)}\Bigg|_{y_1}^{y_2} = \frac{1}{x_2-x_1}$$

$$f_Y(y) = \int_{x_1}^{x_2} \frac{1}{(x_2-x_1)(y_2-y_1)}\,dx = \frac{x}{(x_2-x_1)(y_2-y_1)}\Bigg|_{x_1}^{x_2} = \frac{1}{y_2-y_1}$$



$$E[XY] = \int_{y_1}^{y_2}\int_{x_1}^{x_2} \frac{xy}{(x_2-x_1)(y_2-y_1)}\,dx\,dy = \frac{1}{(x_2-x_1)(y_2-y_1)}\left[\frac{x^2}{2}\Bigg|_{x_1}^{x_2}\right]\left[\frac{y^2}{2}\Bigg|_{y_1}^{y_2}\right] = \frac{(x_1+x_2)(y_1+y_2)}{4}$$

$$F_{X,Y}(x,y) = \int_{x_1}^{x}\int_{y_1}^{y} \frac{1}{(x_2-x_1)(y_2-y_1)}\,dv\,du = \frac{(x-x_1)(y-y_1)}{(x_2-x_1)(y_2-y_1)}$$