ECE 3075A
Random Signals

Lecture 13
Review of Random Variables and Statistics

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

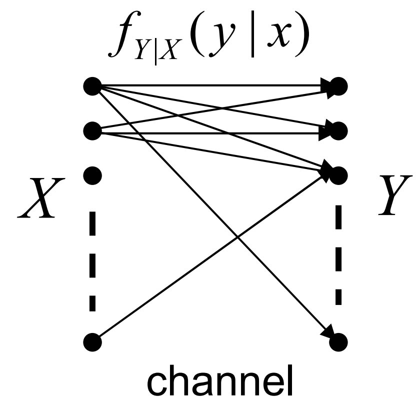
$$\Pr\{\text{HH}\} = 0.3 \times 0.6 = 0.18, \quad \Pr\{\text{HT or TH}\} = \Pr\{\text{HT}\} + \Pr\{\text{TH}\} = 0.3 \times 0.4 + 0.7 \times 0.6 = 0.54, \quad \Pr\{\text{TT}\} = 0.7 \times 0.4 = 0.28$$

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

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

Conditional Density Functions

Let Θ be the phase of a sinusoidal signal that a source sends out to the destination. It is uniformly distributed in the interval $(0, 2\pi)$. During propagation, an independent additive Gaussian noise V, $V \sim N(0, \sigma^2)$, may distort the phase. The signal received at the destination is thus $Y = \Theta + V$. Determine $f_{\Theta|Y} = (\theta \mid y)$.

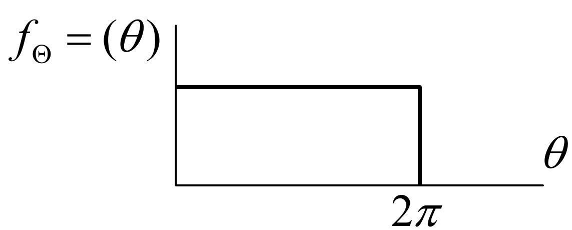
First, recognize that $f_{Y|\Theta}(y \mid \Theta = \theta) = f_V(y - \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-\theta)^2}{2\sigma^2}\right]$

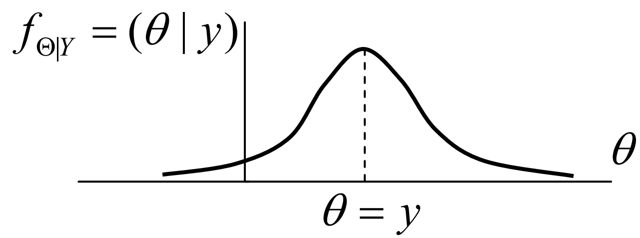
$$f_{Y,\Theta}(y,\theta) = f_{Y|\Theta}(y \mid \theta) f_\Theta(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-\theta)^2}{2\sigma^2}\right] \frac{1}{2\pi}, \quad 0 < \theta < 2\pi$$

A posteriori prob.

$$f_{\Theta|Y}(\theta \mid y) = \frac{f_{Y|\Theta}(y \mid \theta) f_\Theta(\theta)}{\int_0^{2\pi} f_{Y|\Theta}(y \mid \theta) f_\Theta(\theta)} = \frac{\exp\left[-\frac{(y-\theta)^2}{2\sigma^2}\right]}{\int_0^{2\pi} \exp\left[-\frac{(y-\theta)^2}{2\sigma^2}\right] d\theta} = \zeta(y) \exp\left[-\frac{(y-\theta)^2}{2\sigma^2}\right] \quad 0 < \theta < 2\pi$$

$f_{\Theta|Y}(\theta \mid y)$ attains maximum at $\theta = y$ and $\max f_{\Theta|Y}(\theta \mid y) = \zeta(y)$





Statistical Independence

X and Y are statistically independent, then

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

It follows that $E\{X^n Y^k\} = E\{X^n\} E\{Y^k\}$ for all n and k.

When $E\{X^n Y^k\} = E\{X^n\} E\{Y^k\}$ is true for n=k=1, we can only say that X and Y are uncorrelated. To positively prove independence through separation of moments is not advisable.

To prove that X and Y are statistically independent, we need to show that the joint density function is separable, or

$$f_{X,Y}(x \mid y) = f_X(x) \quad \text{or} \quad f_{X,Y}(y \mid x) = f_Y(y)$$

But to prove that X and Y are not independent, we only need to show that at least $E\{X^n Y^k\} \neq E\{X^n\} E\{Y^k\}$ for some n and k. Sometimes this is a lot easier than to find and examine the joint density.

Sum of Random Variables

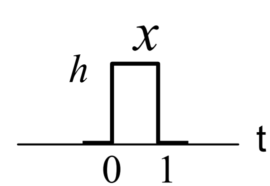
Let $V_i,\ i=1,2,\cdots,L$ be random variables with zero mean and variances $\sigma_{V_1}^2,\sigma_{V_2}^2,\cdots,\sigma_{V_L}^2$, respectively.

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?



Characteristic Function of Sum of R.V.s

- Extend the previous result to sum of many r.v.s

Let $Y = X_1 + X_2 + \cdots X_N$ where $X_1, X_2, \cdots X_N$ are statistically independent r.v.s, with pdf and characteristic functions, $f_{X_i}(x_i)$ and $\Phi_{X_i}(u_i)$, respectively. Their joint characteristic function is thus

$$\Phi_{X_1,X_2,\cdots,X_N}(u_1,u_2,\cdots,u_N) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,X_2,\cdots,X_N}(x_1,x_2,\cdots,x_N)\exp\left[j\sum_{i=1}^{N}u_i x_i\right]dx_1 dx_2\cdots dx_N$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\prod_{i=1}^{N} f_{X_i}(x_i)\right]\exp\left[j\sum_{i=1}^{N}u_i x_i\right]dx_1 dx_2\cdots dx_N = \prod_{i=1}^{N}\int_{-\infty}^{\infty} f_{X_i}(x_i)e^{ju_i x_i}dx_i = \prod_{i=1}^{N}\Phi_{X_i}(u_i)$$

Now, $$\Phi_Y(v) = E[e^{jvY}] = E\left[\exp\left(j\sum_{i=1}^{N}vX_i\right)\right] = \Phi_{X_1,X_2,\cdots,X_N}(v,v,\cdots,v) = \prod_{i=1}^{N}\Phi_{X_i}(v)$$

and $$f_Y(y) = (2\pi)^{-1}\int_{-\infty}^{\infty}\prod_{i=1}^{N}\Phi_{X_i}(v)\,e^{-jvy}dv$$

$\Phi_Y(v)$ is equal to the joint characteristic function along the line $u_1 = u_2 = \cdots = u_N = v$

$$f_Y(y) = (2\pi)^{-1}\int_{-\infty}^{\infty}\left[\Phi_X(v)\right]^N e^{-jvy}dv$$ if X_i are i.i.d. and $\Phi_{X_i}(v) = \Phi_X(v)$

Characteristic Functions

$$Y = X_1 + X_2 + \cdots + X_N \quad \text{sum of independent r.v.s}$$

$$\Phi_Y(v) = E[e^{jvY}] = E\left[\exp\left(j\sum_{i=1}^{N} vX_i\right)\right] = \Phi_{X_1, X_2, \cdots, X_N}(v, v, \cdots, v) = \prod_{i=1}^{N} \Phi_{X_i}(v)$$

$$\frac{d}{dv}\Phi_Y(v) = \frac{d}{dv}\prod_{i=1}^{N}\Phi_{X_i}(v) = \sum_{j=1}^{N}\Phi'_{X_j}(v)\prod_{i=1, i\neq j}^{N}\Phi_{X_i}(v)$$

$$\Phi_X(u)\Big|_{u=0} = E[e^{j0X}] = 1$$

$$j\overline{Y} = \frac{d}{dv}\Phi_Y(v)\Big|_{v=0} = \sum_{j=1}^{N}\Phi'_{X_j}(v)\prod_{i=1, i\neq j}^{N}\Phi_{X_i}(v) = \sum_{j=1}^{n}\Phi'_{X_j}(v)\Big|_{v=0} = \sum_{j=1}^{N} j\overline{X}_j$$

That is, mean of sum is sum of mean.

$$\frac{d^2}{dv^2}\Phi_Y(v) = \frac{d}{dv}\sum_{j=1}^{n}\Phi'_{X_j}(v)\prod_{i=1, i\neq j}^{N}\Phi_{X_i}(v) = \sum_{j=1}^{n}\left\{\Phi''_{X_j}(v)\prod_{i=1, i\neq j}^{N}\Phi_{X_i}(v) + \Phi'_{X_j}(v)\sum_{k=1, k\neq j}^{N}\left(\Phi'_{X_k}(v)\prod_{i=1, i\neq j, i\neq k}^{N}\Phi_{X_i}(v)\right)\right\}$$

$$\sum_{j=1}^{n}\left\{\Phi''_{X_j}(v)\prod_{i=1, i\neq i}^{N}\Phi_{X_i}(v) + \Phi'_{X_j}(v)\sum_{k=1}^{N}\left(\Phi'_{X_k}(v)\prod_{i=1, i\neq i, i\neq k}^{N}\Phi_{X_i}(v)\right)\right\}\Bigg|_{v=0} = \sum_{j=1}^{n}\left\{-\overline{X_j^2} - \overline{X}_j\sum_{k=1, k\neq j}^{N}\overline{X}_k\right\} = -\overline{Y^2}$$

$$\sigma_Y^2 = \overline{Y^2} - (\overline{Y})^2 = \sum_{j=1}^{N}\left\{\overline{X_j^2} + \overline{X}_j\sum_{k=1, k\neq j}^{N}\overline{X}_k\right\} - \left\{\sum_{j=1}^{N}\overline{X}_j\right\}^2 = \sum_{j=1}^{N}\left\{\overline{X_j^2} + (\overline{X}_j)^2\right\} = \sum_{j=1}^{N}\sigma_j^2$$

That is, variance of sum is sum of variances.

Joint Characteristic Function & Joint Moments

$$\Phi_{X,Y}(u,v) = E\left[e^{j(uX+vY)}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)e^{j(uX+vY)}dx\,dy$$

$$f_{X,Y}(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi_{X,Y}(u,v)e^{j(uX+vY)}dx\,dy$$

Note that $\Phi_X(u) = \Phi_{X,Y}(u,0)$ and $\Phi_Y(v) = \Phi_{X,Y}(0,v)$. ⟸ Why?

Use characteristic function to find joint moments:

Recall $\frac{d}{du}\Phi(u)\Big|_{u=0} = \int_{-\infty}^{\infty}\left(\frac{d}{du}e^{jux}\right)f_X(x)\,dx\Big|_{u=0} = \int_{-\infty}^{\infty} jxe^{jux} f_X(x)\,dx\Big|_{u=0} = j\overline{X}$

Similarly, $E[XY] = \overline{XY} = -\left[\frac{\partial^2\Phi_{X,Y}(u,v)}{\partial u\partial v}\right]_{u=v=0}$

Joint moments:

$$E[X^nY^k] = \overline{X^nY^k} = \frac{1}{j^{n+k}}\left[\frac{\partial^{n+k}\Phi_{X,Y}(u,v)}{\partial u^n\partial v^k}\right]_{u=v=0}$$

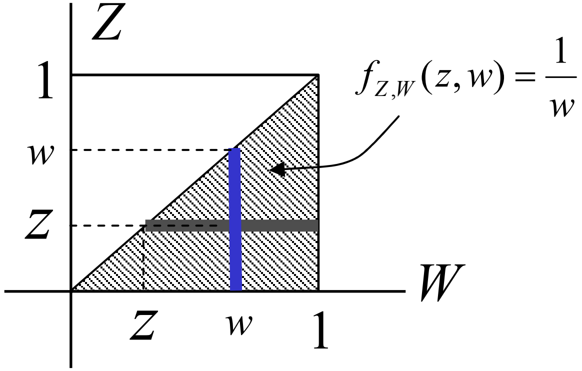
Transformation of Multiple R.V. - Example

Two random variables X and Y have a joint probability density function of the form

$$f_{X,Y}(x,y) = \begin{cases} 1, & 0 < x, y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find the pdf of $Z = XY$.

Use auxiliary function $W = X$



$$Z = XY, \quad W = X \;\Rightarrow\; X = W, \quad Y = Z/W$$

$$J = \begin{vmatrix} \frac{\partial X}{\partial Z} & \frac{\partial X}{\partial W} \\ \frac{\partial Y}{\partial Z} & \frac{\partial Y}{\partial W} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ w^{-1} & -zw^{-2} \end{vmatrix} = -\frac{1}{w}$$

$$f_{Z,W}(z,w) = \frac{1}{|w|} f_{X,Y}\left(w, \frac{z}{w}\right) = \frac{1}{w} \qquad 0 < w < 1 \text{ and } 0 < z < w$$

$$f_Z(z) = \int_z^1 \frac{1}{w} dw = \ln w \Big|_{w=z}^{1} = -\ln(z), \quad 0 < z < w$$

Integration along (black line)

$$f_W(w) = \int_0^w \frac{1}{w} dz = \frac{z}{w}\Big|_{z=0}^{w} = 1, \quad 0 < w < 1 \quad \Leftrightarrow \quad f_X(x) = 1, \quad 0 < x < 1$$

Integration along (blue line)

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 $h(\bullet)$ and y_i is the output, observed with noise, $y_i = h(x_i) + v_i$. With linear regression, $y = g(x) = a + bx$.

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

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

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

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

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