Universidad Carlos III de Madrid

OPEN COURSEWARE CONSORTIUM

Teoría de la Comunicación

Grado en Ingeniería en Tecnologías de Telecomunicación

CAPÍTULO 5

LÍMITES FUNDAMENTALES EN LOS SISTEMAS DE COMUNICACIONES DIGITALES

Marcelino Lázaro

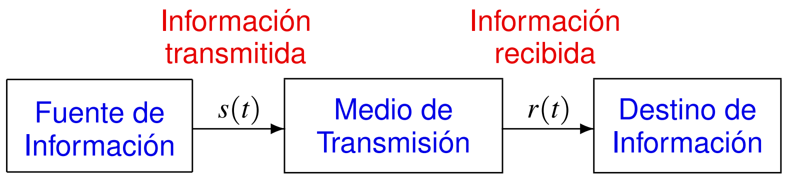
Departamento de Teoría de la Señal y Comunicaciones
Universidad Carlos III de Madrid

Índice de contenidos

- Modelado de fuentes de información
- Modelos probabilísticos de canal
- Medidas cuantitativas de información
- Capacidad de canal
  - Capacidad de un canal digital
  - Capacidad de un canal gausiano
- Límites en un sistema digital de comunicaciones

Definición de un sistema de comunicaciones

- Finalidad de un sistema de comunicaciones:
  - *Transmisión de información*



- Cuantificación de la información
  - Medidas cuantitativas de información
  - Análisis de un sistema de comunicaciones
    - Cantidad de información generada
    - Límites fundamentales en la transmisión de información
- Organización del capítulo:
  - Modelos (probabilísticos) para las fuentes de información
  - Modelos (probabilísticos) del sistema (de canal)
  - Medidas cuantitativas de información
  - Límites fundamentales en los sistemas de comunicación

Modelado de fuentes de información

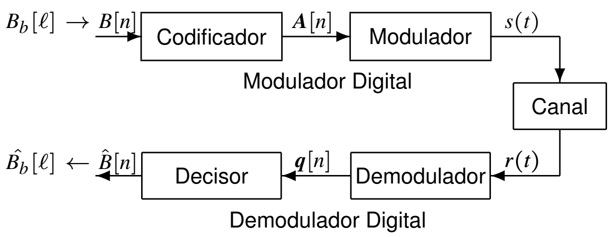
- Salida de la fuente: flujo de información
  - En tiempo continuo: señal $x(t)$
  - En tiempo discreto: señal $x[n]$
- Modelo de la salida de la fuente (información)
  - Proceso aleatorio, $X(t)$, ó $X[n]$
- Modelo para fuentes en tiempo continuo (analógicas)
  - Proceso aleatorio en tiempo continuo $X(t)$
    - Caracterízación: densidad espectral de potencia $S_X(j\omega)$
      - Normalmente son procesos limitados en banda
      - Refleja el comportamiento espectral medio de la fuente
- Modelo para fuente en tiempo discreto
  - Proceso aleatorio en tiempo discreto $X[n]$
  - Tipos de alfabeto de la fuente
    - Continuo (e.g., señales muestreadas)
    - Discreto (fuentes digitales)
      - Modelo más simple: fuente discreta sin memoria

Fuente discreta sin memoria

DMS : *Discrete Memoryless Source*

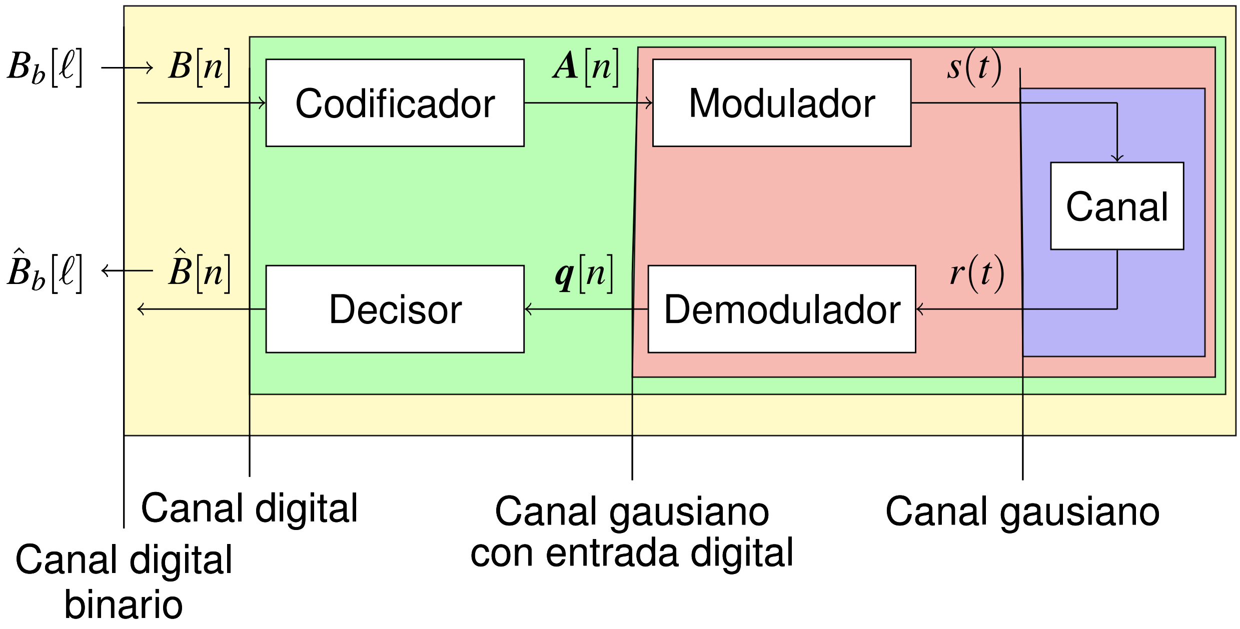
- Proceso aleatorio en tiempo discreto $X[n]$
  - Alfabeto discreto de M_X valores
  - Sin memoria: $\{X[n]\}$ son variables aleatorias independientes e idénticamente distribuidas (i.i.d.)
- Descripción del proceso (caracterización estadística)
  - Variable aleatoria X (al ser $X[n]$ i.i.d., la estadística es la misma para todo n)
    - ★ Alfabeto $\mathcal{A}_X = \{x_0, x_1, \cdots, x_{M_X-1}\}$
    - ★ Probabilidades $p_X(x_i) = P(X = x_i)$ para $i = 0, 1, \cdots, M_X - 1$
- Ejemplo: fuente binaria simétrica
  BSS: *Binary Symmetric Source*
  - Alfabeto $\mathcal{A}_X = \{x_0, x_1\}$, típicamente $x_0 = 0$, $x_1 = 1$
  - Probabilidades $p_X(x_0) = p_X(x_1) = \frac{1}{2}$

Modelos probabilísticos de canal



- Modelos probabilísticos de canal
  - Establecen la relación probabilística entre la información recibida y la transmitida en distintos puntos de este modelo general del sistema de comunicaciones
    - ★ Distintos niveles de abstracción en la definición de *canal*
    - ★ Modelo probabilístico: Entrada $\boxed{X}$, salida $\boxed{Y}$, distribución $\boxed{f_{Y|X}(y|x)}$
- Modelos definidos
  - Canal gaussiano
    - ★ Representa el canal físico: $Y \equiv r(t) \mid X \equiv s(t)$
  - Canal gaussiano con entrada digital
    - ★ Representa el canal discreto equivalente: $Y \equiv \boldsymbol{q}[n] \mid X \equiv \boldsymbol{A}[n]$
  - Canal digital
    - ★ Representa la transmisión de símbolos: $Y \equiv \hat{B}[n] \mid X \equiv B[n]$ (ó $\hat{\boldsymbol{A}}[n] \mid \boldsymbol{A}[n]$)
  - Canal digital binario
    - ★ Representa la transmisión de bits: $Y \equiv \hat{B}_b[\ell] \mid X \equiv B_b[\ell]$

Modelos probabilísticos de canal - Representación

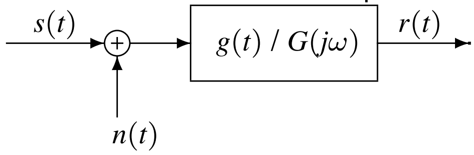


Canal gaussiano

- Relación entrada / salida
  - Entrada: $X \equiv s(t)$, para un instante dado t
  - Salida: $Y \equiv r(t)$, para el mismo instante dado t
- Modelo canal gaussiano

$$r(t) = s(t) + n(t)$$

  - $n(t)$: proceso aleatorio blanco y gaussiano con $S_n(j\omega) = \frac{N_0}{2}$, $R_n(\tau) = \frac{N_0}{2} \cdot \delta(\tau)$
- Limitación de la potencia de ruido - Filtrado a la entrada del receptor
  - Señal de ancho de banda B Hz ($W = 2\pi B$ rad/s)
    - ★ Filtro $g(t)$ ideal de ancho de banda B Hz: potencia de ruido $N_0 \cdot B$



- Distribución de los valores de la señal $r(t)$ dada la señal $s(t)$ en un cierto instante t (distribución $Y|X$ cuando $Y \equiv r(t)$ y $X \equiv s(t)$)
  - Distribución de Y cuando $X = x = s(t)$
  - Gaussiana, de media $s(t)$ en el instante dado y varianza $\sigma^2 = N_0 \cdot B$

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}$$

Canal gaussiano con entrada digital

- Es el equivalente al canal discreto equivalente
- Relación entrada / salida
  - Entrada: $\boldsymbol{X} \equiv \boldsymbol{A}[n]$, para un instante dado n
    - ★ Vector de N variables aleatorias discretas
    - ★ Si $\boldsymbol{A}[n] = \boldsymbol{a}_i \rightarrow \boldsymbol{X} = \boldsymbol{x}_i$
  - Salida: $\boldsymbol{Y} \equiv \boldsymbol{q}[n]$, para el mismo instante dado n
    - ★ Vector de N variables aleatorias continuas
- Modelo del canal gaussiano con entrada digital

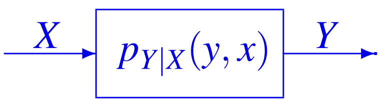
$$\boldsymbol{q} = \boldsymbol{A} + \boldsymbol{z}$$

  - Distribución de los N elementos del vector de ruido
    - ★ V.A.'s Independientes, gaussianas, media nula, varianza $N_0/2$
- Distribución de la salida dada la entrada
  - Gaussiana de media el símbolo transmitido ($\boldsymbol{x}_i \equiv \boldsymbol{a}_i$)

$$f_{\boldsymbol{q}|\boldsymbol{A}}(\boldsymbol{q}|\boldsymbol{a}_i) = \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{||\boldsymbol{q}-\boldsymbol{a}_i||^2}{N_0}} \rightarrow f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}_i) = \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{||\boldsymbol{y}-\boldsymbol{x}_i||^2}{N_0}}$$

Canal digital

- Relación entrada / salida
  - Entrada: $X \equiv B[n]$, para un instante dado n
  - Salida: $Y \equiv \hat{B}[n]$, para el mismo instante dado n
    - ★ Alfabeto de X e Y
      - M elementos del alfabeto de $B[n]$: $\{x_i = y_i \equiv b_i\}_{i=0}^{M-1}$
- Modelo probabilístico: canal discreto sin memoria
  DMC: *Discrete Memoryless Channel*



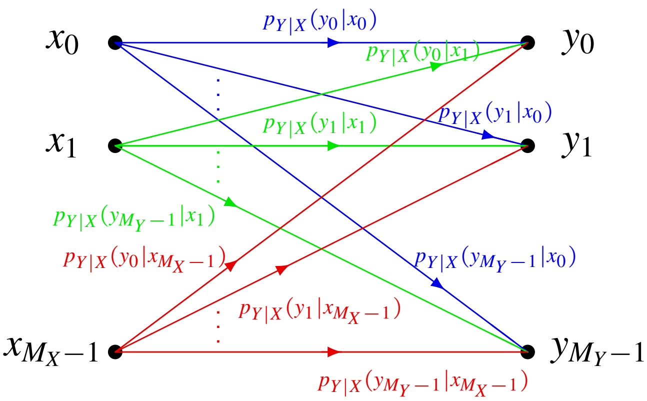
  - Definición estadística del canal discreto sin memoria
    1. *Alfabeto de entrada*
$$\mathcal{A}_X = \{x_i\}, \; i = 0, \cdots, M_X - 1$$
    2. *Alfabeto de salida*
$$\mathcal{A}_Y = \{y_j\}, \; j = 0, \cdots, M_Y - 1$$
    3. *Conjunto de probabilidades condicionales (de transición)*
$$p_{Y|X}(y_j|x_i), \; \forall i, j$$

Representación de las probabilidades de transición

- Matriz de canal

$$\boldsymbol{P} = \begin{bmatrix} p_{Y|X}(y_0|x_0) & p_{Y|X}(y_1|x_0) & \cdots & p_{Y|X}(y_{M_Y-1}|x_0) \\ p_{Y|X}(y_0|x_1) & p_{Y|X}(y_1|x_1) & \cdots & p_{Y|X}(y_{M_Y-1}|x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_{Y|X}(y_0|x_{M_X-1}) & p_{Y|X}(y_1|x_{M_X-1}) & \cdots & p_{Y|X}(y_{M_Y-1}|x_{M_X-1}) \end{bmatrix}$$

  - Elementos de una fila suman 1
- Diagrama de flechas



  - Flechas que salen de un mismo nodo suman 1

Aplicación del DMC al canal digital

- Alfabetos de X e Y
  - M símbolos del alfabeto de $B[n]$, $\{b_0, b_1, \cdots, b_{M-1}\}$

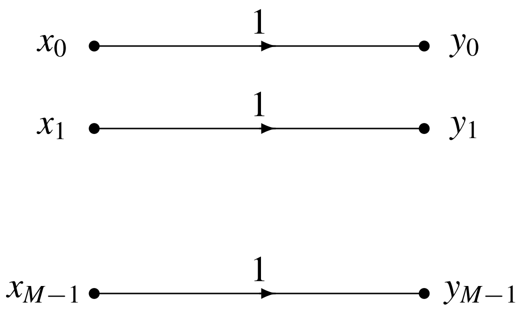
$$x_i \equiv b_i,\; y_j \equiv b_j,\; M_X = M_Y = M,\; i,j \in \{0, 1, \cdots M-1\}$$

- Probabilidades de transición

$$p_{Y|X}(y_j|x_i) \equiv p_{\hat{B}|B}(b_j|b_i) = p_{\hat{\boldsymbol{A}}|\boldsymbol{A}}(\boldsymbol{a}_j|\boldsymbol{a}_i)$$

  - Valores ideales: matriz de transición / diagrama de flechas

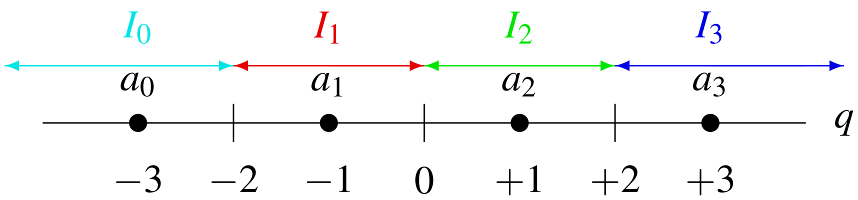
$$\boldsymbol{P} = \boldsymbol{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & 1 & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$



Cálculo de las probabilidades de transición - Ejemplo

- $M = 4$, símbolos equiprobables $p_A(\boldsymbol{a}_i) = \frac{1}{4}$
  - Constelación: $\boldsymbol{a}_0 = -3,\ \boldsymbol{a}_1 = -1,\ \boldsymbol{a}_2 = +1,\ \boldsymbol{a}_3 = +3$
  - Regiones de decisión: umbrales $q_{u1} = -2,\ q_{u2} = 0,\ q_{u3} = +2$

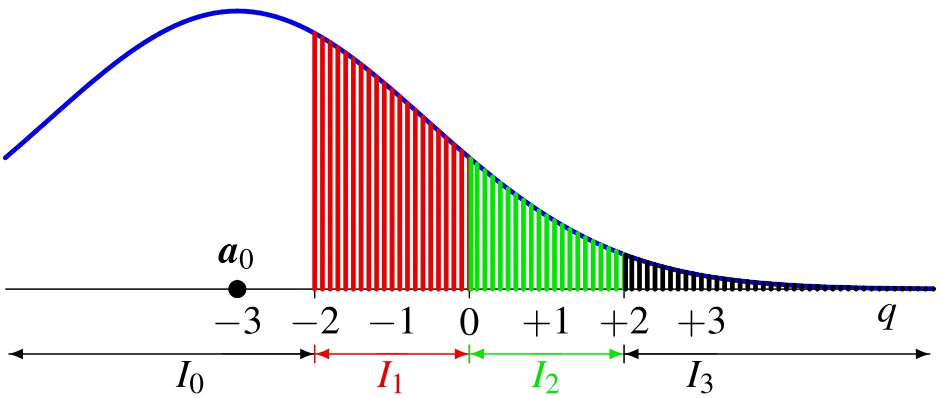
$$I_0 = (-\infty, -2],\ I_1 = (-2, 0],\ I_2 = (0, +2],\ I_3 = (+2, +\infty)$$



- Probabilidades de transición (matriz de canal)

$$\boldsymbol{P} = \begin{bmatrix} 1 - Q\left(\frac{1}{\sqrt{N_o/2}}\right) & Q\left(\frac{1}{\sqrt{N_o/2}}\right) - Q\left(\frac{3}{\sqrt{N_o/2}}\right) & Q\left(\frac{3}{\sqrt{N_o/2}}\right) - Q\left(\frac{5}{\sqrt{N_o/2}}\right) & Q\left(\frac{5}{\sqrt{N_o/2}}\right) \\ Q\left(\frac{1}{\sqrt{N_o/2}}\right) & 1 - 2Q\left(\frac{1}{\sqrt{N_o/2}}\right) & Q\left(\frac{1}{\sqrt{N_o/2}}\right) - Q\left(\frac{3}{\sqrt{N_o/2}}\right) & Q\left(\frac{3}{\sqrt{N_o/2}}\right) \\ Q\left(\frac{3}{\sqrt{N_o/2}}\right) & Q\left(\frac{1}{\sqrt{N_o/2}}\right) - Q\left(\frac{3}{\sqrt{N_o/2}}\right) & 1 - 2Q\left(\frac{1}{\sqrt{N_o/2}}\right) & Q\left(\frac{1}{\sqrt{N_o/2}}\right) \\ Q\left(\frac{5}{\sqrt{N_o/2}}\right) & Q\left(\frac{3}{\sqrt{N_o/2}}\right) - Q\left(\frac{5}{\sqrt{N_o/2}}\right) & Q\left(\frac{1}{\sqrt{N_o/2}}\right) - Q\left(\frac{3}{\sqrt{N_o/2}}\right) & 1 - Q\left(\frac{1}{\sqrt{N_o/2}}\right) \end{bmatrix}$$

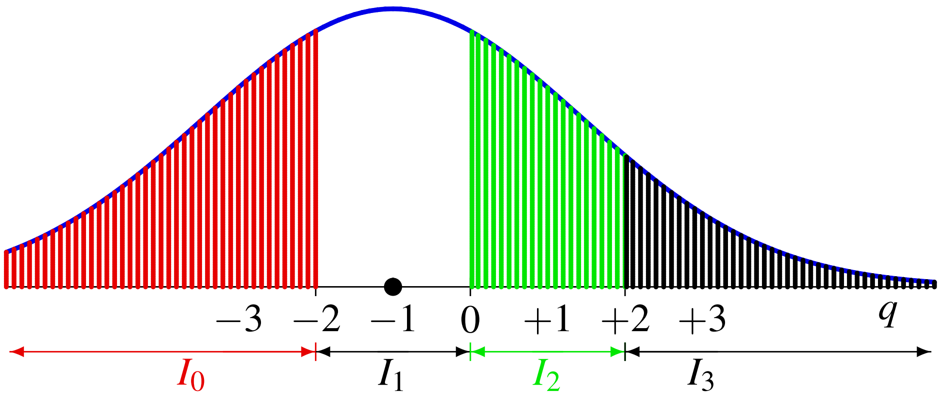
Elementos de la primera fila: $p_{Y|X}(y_j|x_0),\ \forall j$



- Distribución $f_{\boldsymbol{q}|\boldsymbol{A}}(\boldsymbol{q}|\boldsymbol{a}_0)$: gaussiana de media $\boldsymbol{a}_0$ y varianza $N_0/2$

$$\begin{aligned} p_{Y|X}(y_0|x_0) &= 1 - P_{e|\boldsymbol{a}_0} = 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right) \\ p_{Y|X}(y_1|x_0) &= P_{e|\boldsymbol{a}_0 \to \boldsymbol{a}_1} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) \\ p_{Y|X}(y_2|x_0) &= P_{e|\boldsymbol{a}_0 \to \boldsymbol{a}_2} = Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) \\ p_{Y|X}(y_3|x_0) &= P_{e|\boldsymbol{a}_0 \to \boldsymbol{a}_3} = Q\left(\frac{5}{\sqrt{N_0/2}}\right) \end{aligned}$$

Elementos de la segunda fila: $p_{Y|X}(y_j|x_1)$, $\forall j$



- Distribución $f_{\boldsymbol{q}|\boldsymbol{A}}(\boldsymbol{q}|\boldsymbol{a}_1)$: gaussiana de media $\boldsymbol{a}_1$ y varianza $N_0/2$

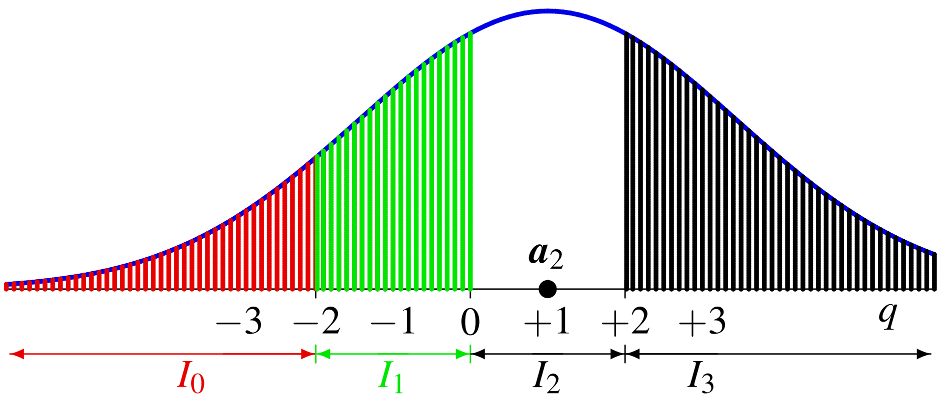
$$p_{Y|X}(y_0|x_1) = P_{e|\boldsymbol{a}_1 \to \boldsymbol{a}_0} = Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_1|x_1) = 1 - P_{e|\boldsymbol{a}_1} = 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_2|x_1) = P_{e|\boldsymbol{a}_1 \to \boldsymbol{a}_2} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_3|x_1) = P_{e|\boldsymbol{a}_1 \to \boldsymbol{a}_3} = Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

Elementos de la tercera fila: $p_{Y|X}(y_j|x_2)$, $\forall j$



- Distribución $f_{\boldsymbol{q}|\boldsymbol{A}}(\boldsymbol{q}|\boldsymbol{a}_2)$: gaussiana de media $\boldsymbol{a}_2$ y varianza $N_0/2$

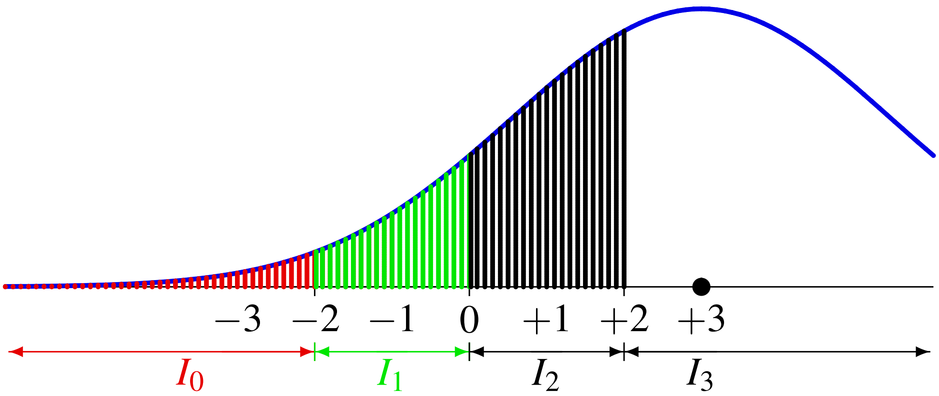
$$p_{Y|X}(y_0|x_2) = P_{e|\boldsymbol{a}_2 \to \boldsymbol{a}_0} = Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_1|x_2) = P_{e|\boldsymbol{a}_2 \to \boldsymbol{a}_1} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_2|x_2) = 1 - P_{e|\boldsymbol{a}_2} = 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_3|x_2) = P_{e|\boldsymbol{a}_2 \to \boldsymbol{a}_3} = Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

Elementos de la cuarta fila: $p_{Y|X}(y_j|x_3)$, $\forall j$



- Distribución $f_{q|A}(q|a_3)$: gaussiana de media a_3 y varianza $N_0/2$

$$p_{Y|X}(y_0|x_3) = P_{e|a_3 \to a_0} = Q\left(\frac{5}{\sqrt{N_0/2}}\right)$$

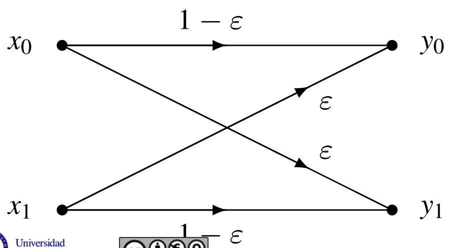
$$p_{Y|X}(y_1|x_3) = P_{e|a_3 \to a_1} = Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_2|x_3) = P_{e|a_3 \to a_2} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_3|x_3) = 1 - P_{e|a_3} = 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

Canal digital binario

- Relación entrada / salida
  - Entrada: $X \equiv B_b[\ell]$, para un instante dado ℓ
  - Salida: $Y \equiv \hat{B}_b[\ell]$, para el mismo instante dado n
    - Alfabeto de X e Y: Bits $x_0 = y_0 \equiv 0$, $x_1 = y_1 \equiv 1$
- Modelo probabilístico
  - Particularización del DMC para $M = 2$
  - Probabilidades condicionales $p_{Y|X}(y_j|x_i)$
    - Probabilidades de acierto de bit ($j = i$)
    - Probabilidades de error de bit ($j \neq i$)
  - Habitualmente, se asume que la probabilidad de error de bit es igual para los dos bits $(0, 1)$
    - Canal binario simétrico (BSC:*Binary Symmetric Channel*)
    - Probabilidad de error de bit (BER): $BER = \varepsilon$



$$\mathbf{P} = \begin{bmatrix} p_{Y|X}(y_0|x_0) & p_{Y|X}(y_1|x_0) \\ p_{Y|X}(y_0|x_1) & p_{Y|X}(y_1|x_1) \end{bmatrix} = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix}$$

MEDIDAS CUANTITATIVAS

DE

INFORMACIÓN

Autoinformación

- Medida de la información contenida en un suceso aislado de una variable aleatoria discreta ($X = x_i$)
- Requerimientos para esta medida
  1. $I_X(x_i)$ depende de la probabilidad del suceso
  $$I_X(x_i) = f(p_X(x_i))$$
  2. Debe ser una función decreciente de la probabilidad
  $$p_X(x_i) > p_X(x_j) \rightarrow I_X(x_i) < I_X(x_j)$$
  3. Debe ser una función continua de la probabilidad
     - Sucesos con probabilidades similares tienen una información similar
  4. Para sucesos independientes $X = x_i$, $Y = y_j$ ($p_{X,Y}(x_i, y_j) = p_X(x_i) \cdot p_Y(y_j)$)
  $$I_{X,Y}(x_i, y_j) = I_X(x_i) + I_Y(y_j)$$
- Función que cumple estas propiedades - Autoinformación
$$\boxed{I_X(x_i) = -\log_b(p_X(x_i))}$$
  - La base del logaritmo define las unidades de la medida
    - Base 2 : bits
    - Base e (logaritmo natural o neperiano $\ln$): nats
    - NOTA: Relación $\log_b(x) = \ln(x)/\ln(b)$

Entropía

- Medida de la incertidumbre contenida en una variable aleatoria (discreta)
  - Promedio de la autoinformación de cada suceso
    - ★ Alfabeto: $\mathcal{A}_X = \{x_0, x_1, \cdots, x_{M_X-1}\}$
    - ★ Probabilidades: $\{p_X(x_0), p_X(x_1), \cdots, p_X(x_{M_X-1})\}$

$$H(X) = -\sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log p_X(x_i) = \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log\left(\frac{1}{p_X(x_i)}\right)$$

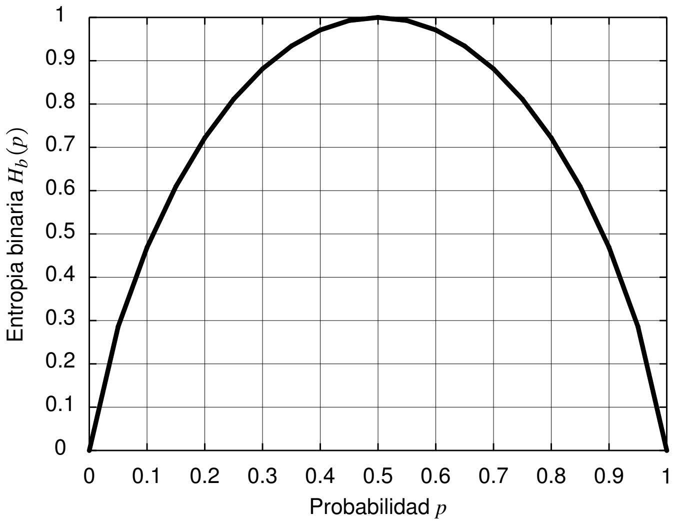
NOTA: Por convención: $0 \times \log 0 = 0$

  - Unidades: bits/símbolo ó nats/símbolo
- Valores límite de la entropía de variables aleatorias discretas
  1. $H(X) \geq 0$
     Ya que $0 \leq p_X(x_i) \leq 1$ y, consecuentemente, $\log(1/p_X(x_i)) \geq 0$
     - ★ $H(X) = 0$ cuando no hay incertidumbre en X
       $p_X(x_i) = 1, p_X(x_j) = 0 \; \forall j \neq i$
  2. $H(X) \leq \log(M_X)$
     La igualdad se produce únicamente si los símbolos son equiprobables, $p_X(x_i) = \frac{1}{M_X}$
     - ★ Máxima situación de incertidumbre

Ejemplo - Entropía binaria

- Variable aleatoria binaria
  - Alfabeto: $\{x_0, x_1\}$
  - Probabilidades: $\{p_X(x_0) = p, p_X(x_1) = 1 - p\}$

$$H(X) = -p \cdot \log(p) - (1-p) \cdot \log(1-p)$$
$$= p \cdot \log\left(\frac{1}{p}\right) + (1-p) \cdot \log\left(\frac{1}{1-p}\right) \equiv H_b(p)$$



Entropía conjunta

- Medida de la información conjunta de dos (o más) variables aleatorias

$$H(X,Y) = \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j) \cdot \log \left(\frac{1}{p_{X,Y}(x_i,y_j)} \right)$$

- Variables aleatorias independientes
  - Probabilidad conjunta: $p_{X,Y}(x_i,y_j) = p_X(x_i) \cdot p_Y(y_j)$
  - Entropía conjunta

$$\begin{aligned}
H(X,Y) &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_X(x_i) \cdot p_Y(y_j) \cdot \log \frac{1}{p_X(x_i) \cdot p_Y(y_j)} \\
&= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_X(x_i) \cdot p_Y(y_j) \cdot \log \frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_X(x_i) \cdot p_Y(y_j) \cdot \log \frac{1}{p_Y(y_j)} \\
&= \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log \frac{1}{p_X(x_i)} + \sum_{j=0}^{M_Y-1} p_Y(y_j) \cdot \log \frac{1}{p_Y(y_j)} \\
&= H(X) + H(Y)
\end{aligned}$$

Entropía condicional

- Mide la incertidumbre en una variable aleatoria cuando se conoce el valor de la otra
  - Promedio de $H(X|Y=y_j)$ para cada valor del alfabeto de Y

$$\begin{aligned}
H(X|Y) &= \sum_{j=0}^{M_Y-1} p_Y(y_j) \cdot H(X|Y=y_j) \\
&= \sum_{j=0}^{M_Y-1} p_Y(y_j) \sum_{i=0}^{M_X-1} p_{X|Y}(x_i|y_j) \cdot \log \frac{1}{p_{X|Y}(x_i|y_j)} \\
&= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j) \cdot \log \frac{1}{p_{X|Y}(x_i|y_j)}
\end{aligned}$$

  - Para variables aleatorias independientes

$$H(X|Y) = H(X)$$

Relación entre entropía conjunta y condicional

$$
\begin{aligned}
H(X,Y) &= \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_{X,Y}(x_i,y_j)}\\
&= \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_X(x_i)\cdot p_{Y|X}(y_j|x_i)}\\
&= \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_{Y|X}(y_j|x_i)}\\
&= \sum_{i=0}^{M_X-1} p_X(x_i)\cdot\log\frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_{Y|X}(y_j|x_i)}\\
&= H(X)+H(Y|X)
\end{aligned}
$$

$$\boxed{H(X,Y)=H(X)+H(Y|X)=H(Y)+H(X|Y)}$$

Información mutua

- Mide la información que aporta una variable aleatoria X sobre el conocimiento de otra variable aleatoria Y

$$I(X,Y)=\sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{p_{X,Y}(x_i,y_j)}{p_X(x_i)\cdot p_Y(y_j)}$$

- Propiedades
  1. $I(X,Y)=I(Y,X)\geq 0$
     La igualdad se cumple en el caso de que X e Y sean independientes
  2. $I(X,Y)\leq \text{mín}(H(X),H(Y))$
  3. Se puede definir información mutua condicional
     $$I(X,Y|Z)=\sum_{i=0}^{M_Z-1} p_Z(z_i)\cdot I(X,Y|Z=z_i)$$
     $$I(X,Y|Z)=H(X|Z)-H(X|Y,Z)$$
  4. La regla de la cadena para la información mutua es
     $$I((X,Y),Z)=I(X,Z)+I(Y,Z|X)$$
     $$I((X_1,X_2,\cdots,X_N),Y)=I(X_1,Y)+I(X_2,Y|X_1)+\cdots+I(X_N,Y|X_1,\cdots,X_{N-1})$$
  5. A partir de la definición de información mutua se obtiene la definición de entropía
     $$I(X,X)=H(X)$$

Relaciones información mutua y entropía

$$
\begin{aligned}
I(X,Y) &= \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{p_{X,Y}(x_i,y_j)}{p_X(x_i)\cdot p_Y(y_j)}\\
&= \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{p_{X|Y}(x_i,y_j)}{p_X(x_i)}\\
&= \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log(p_{X|Y}(x_i,y_j))\\
&= \sum_{i=0}^{M_X-1} p_X(x_i)\cdot\log\frac{1}{p_X(x_i)} - \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{X,Y}(x_i,y_j)\cdot\log\frac{1}{p_{X|Y}(x_i,y_j)} = H(X)-H(X|Y)
\end{aligned}
$$

$$\boxed{I(X,Y)=H(X)-H(X|Y)=H(Y)-H(Y|X)=H(X)+H(Y)-H(X,Y)}$$

Relaciones información mutua y entropía (II)

$$\boxed{I(X,Y)=H(X)-H(X|Y)=H(Y)-H(Y|X)=H(X)+H(Y)-H(X,Y)}$$

- Representación en un diagrama de Venn
  - Entropías e información mutua representadas por áreas

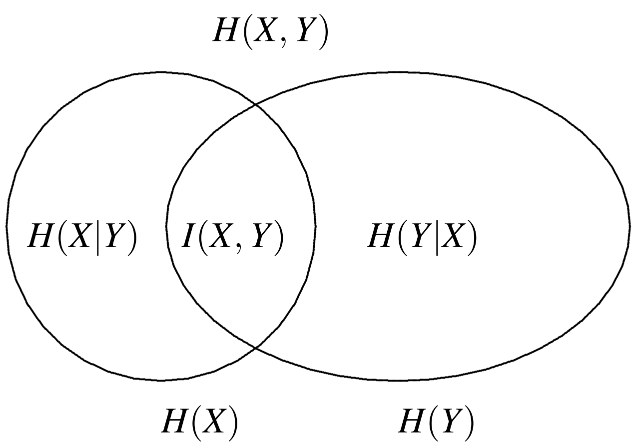
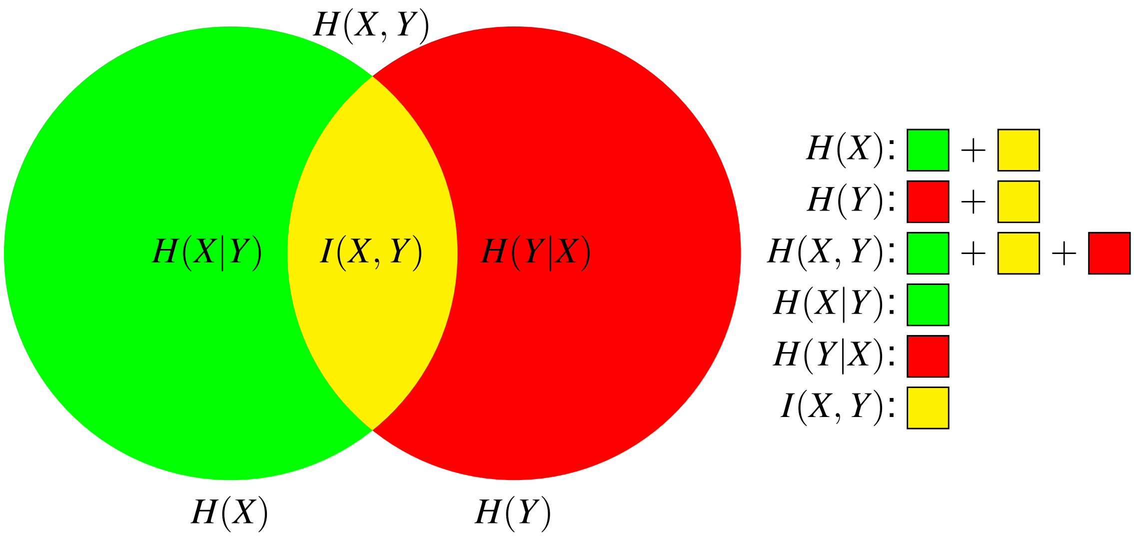


Diagrama de Venn - Entropías e información mutua



Entropía diferencial

- Extensión de las definiciones de entropía a variables aleatorias continuas

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \cdot \log \frac{1}{f_X(x)} \, dx$$

Definición de *entropía diferencial conjunta*

$$h(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log \frac{1}{f_{X,Y}(x, y)} dx \, dy$$

Lo mismo se hace para la *entropía diferencial condicional*

$$h(X|Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log \frac{1}{f_{X|Y}(x|y)} dx \, dy$$

A menudo se utiliza la definición alternativa pero equivalente

$$h(X|Y) = \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} f_{X|Y}(x|y) \log \frac{1}{f_{X|Y}(x|y)} dx \, dy$$

Entropía diferencia e información mutua - Relaciones

- Definición de la información mutua

$$I(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\cdot \log \frac{f_{X,Y}(x,y)}{f_X(x)\cdot f_Y(y)} dx\, dy$$

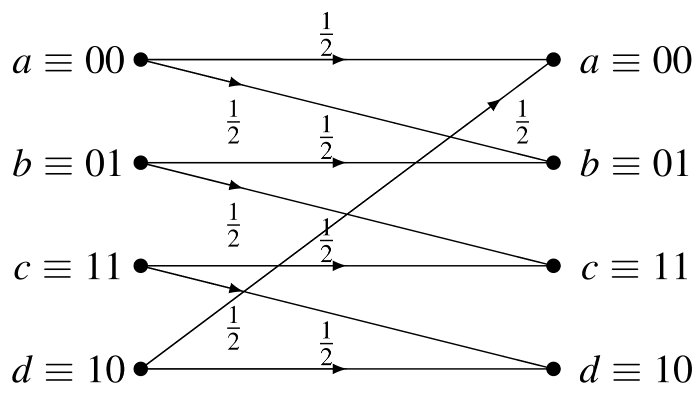
- Se mantienen las mismas relaciones que para variables discretas

$$h(X,Y) = h(X) + h(Y|X) = h(Y) + h(X|Y)$$

$$I(X,Y) = h(Y) - h(Y|X) = h(X) - h(X|Y) = h(X) + h(Y) - h(X,Y)$$

Límites Fundamentales en Sistemas de Comunicaciones

Transmisión fiable sobre canales no fiables - Ejemplo



- 4 símbolos $\equiv$ 2 bits de información por uso del canal
- El canal no es fiable - Se comenten errores
  - Probabilidad de error de símbolo es $P_e = 1/2$
  - Con asignación binaria de Gray $BER = 1/4$
  - Los errores se producen porque dada una salida no es posible identificar de forma unívoca el símbolo transmitido
    - Por ejemplo: si se observa a, esto puede ser porque
      - Se ha transmitido a (sin error)
      - Se ha transmitido d (con error)

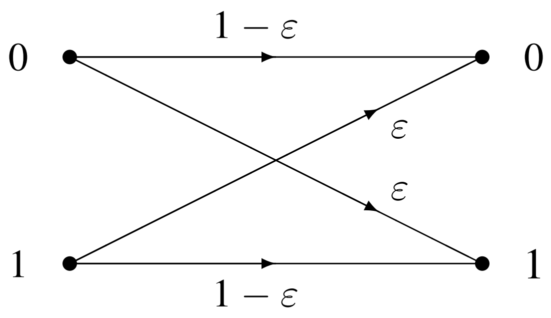
Transmisión fiable sobre canales no fiables - Ejemplo (II)

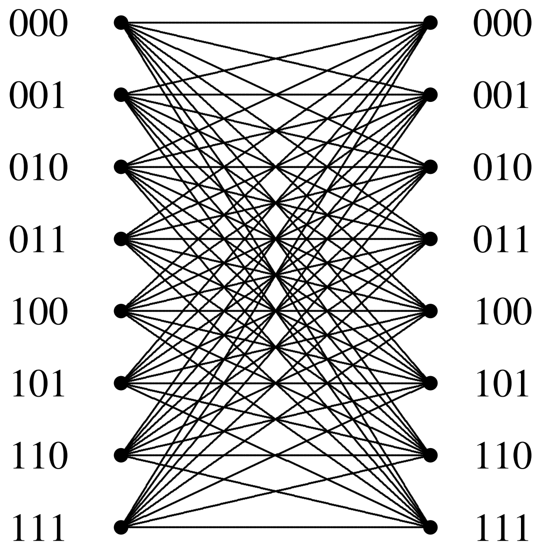
- Opción para transmitir información de forma fiable
  - Transmitir un subconjunto de los símbolos
    - Símbolos que generen salidas “*sin solapamiento*”
  - Ejemplo: transmitir sólo a y c
    - a puede dar como salidas a o b
    - c puede dar como salidas c o d

  Dada una salida no hay incertidumbre en el símbolo transmitido !!!
- Es posible transmitir información sobre este canal con probabilidad de error nula
  - Coste de la transmisión fiable
    - Por cada uso del canal se transmite menos información
      - En este caso: 2 símbolos $\equiv$ 1 bit por uso del canal
- Los canales habituales no permiten esto de forma directa
  - Alternativa: forzar un comportamiento similar - **Codificación**
    - No se busca probabilidad de error nula (sin solapamientos)
    - Se busca poder reducir la probabilidad de error de forma arbitraria
      - Solapamientos con probabilidad arbitrariamente baja

Codificación

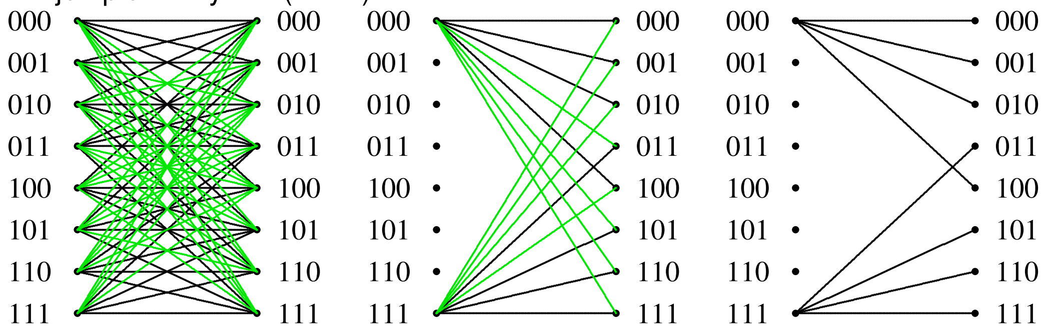
- Se utiliza el canal n veces de forma conjunta
  - Definición de símbolos extendidos: grupo de n símbolos
- Se busca un subconjunto de símbolos (2^k) que produzcan "*bajo solapamiento*" en la salida
  - Se transportan k bits de información por cada n usos del canal
- Ejemplo: canal binario simétrico ($BER = \varepsilon$) con $n = 3$





Codificación (II)

- Situaciones más probables (para ε razonablemente bajo)
  - 0 errores o 1 error sobre 3 bits - 4 ramas/símbolo (en negro)
- Situaciones menos probables
  - 2 errores o 3 errores sobre 3 bits - 4 ramas/símbolo (en verde)
- Subconjunto de 2^k ($k < n$) elementos con "*bajo solapamiento*"
  - Ejemplo: 000 y 111 ($k = 1$)



- Despreciando los enlaces de "*baja probabilidad*" no hay solapamiento
  - Probabilidad de error: $P_e = 3 \times \left[\varepsilon^2 \cdot (1-\varepsilon)\right] + \varepsilon^3$
    - ★ Ejemplo A: $\varepsilon = 0{,}1 \rightarrow P_e = 0{,}028$ | Ejemplo B: $\varepsilon = 0{,}01 \rightarrow P_e = 0{,}000298$
  - Información transmitida: 1 bit (k) de información por cada 3 (n) usos del canal
    - ★ Tasa $R = k/n = 1/3$
- Intuición: aumentando n y k (con k/n constante) se puede reducir más
  - Existe un límite: $\boxed{\text{Capacidad de canal}}$

Codificación de canal para protección frente a errores

- Introducción de redundancia estructurada para reducir la probabilidad de error
  - Bloques de k símbolos de información (conjunto de índices) se transforman en bloques de n bits (paralbras del codigo)
    - Diccionario del código

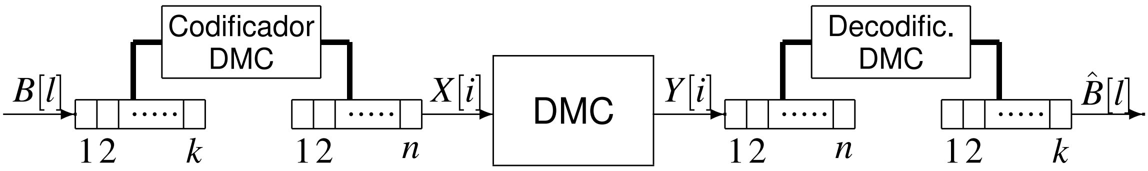
Ejemplo de diccionario del código para dos códigos binarios

Conjunto de índices	Palabras código
0	000
1	111

Código de ejemplo $\mathcal{C}(1,3)$

Conjunto de índices	Palabras código
00	00000
01	10101
10	01110
11	11011

Código de ejemplo $\mathcal{C}(2,5)$

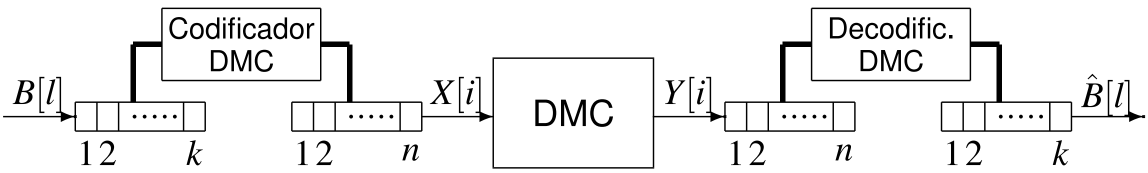


Teorema de codificación de canal

$$C = \max_{p_X(x_i)} I(X,Y)$$

Teorema de codificación de canal (Shannon 1948):

1. Si la tasa de transmisión R es menor que C, entonces para cualquier $\delta > 0$ existe un código con una longitud de bloque n suficientemente larga cuya probabilidad de error es menor que δ
2. Si $R > C$, la probabilidad de error de cualquier código con cualquier longitud de bloque está limitada por un valor no nulo
3. Existen códigos que permiten alcanzar la capacidad del canal $R = C$



Tasa del código: $R = \dfrac{k}{n}$

Capacidad de canal

- Máxima cantidad de información que se puede transmitir de forma fiable a través de un canal de comunicaciones en un sistema digital de comunicaciones
  - En la transmisión se produce una distorsión
    - ★ Potencial pérdida de información
  - Transmisión fiable - Definición
    - ★ Transmisión sin pérdida potencial de información
    - ★ En la práctica: capacidad para reducir la probabilidad de error tanto como sea necesario
  - Capacidad de canal
    - ★ Límite en el número de símbolos con solapamiento arbitrariamente bajo cuando el número de usos del canal tiende a infinito
  - Introducción al concepto de codificación de canal
    - ★ Mecanismo que permite una transmisión fiable
- Se estudiarán los siguientes casos:
  - Canal digital
    - ★ El canal digital binario se considera un caso particular
  - Canal gaussiano

Capacidad de canal a través de la información mutua

- Información mutua entre la entrada y la salida de un DMC

$$I(X,Y) = H(X) - H(X|Y)$$

- Análisis para un BSC con $BER = \varepsilon$ en dos casos extremos
  - Caso óptimo: $\varepsilon = 0$
    - ★ $H(X|Y) = 0 \rightarrow I(X,Y) = H(X)$
  - Peor caso: $\varepsilon = 1/2$
    - ★ $H(X|Y) = H(X) \rightarrow I(X,Y) = 0 \rightarrow X$ e Y independientes
- Se pueden extraer las siguientes conclusiones
  1. La información mutua entre entrada y salida del canal se puede ver como la cantidad de información que pasa de la entrada a la salida cuando el canal es utilizado. En el caso en que la probabilidad de error es nula, pasa toda la información ($I(X,Y) = H(X)$), y en el caso en que la entrada y la salida son estadísticamente independientes se "pierde" toda la información ($I(X,Y) = 0$).
  2. $H(X|Y)$ puede interpretarse como la información que se "pierde" en el canal, y así la información que "atraviesa" el canal, $I(X,Y)$, es igual a la información que hay a la entrada, $H(X)$, menos la que se pierde, $H(X|Y)$. Cuando la probabilidad de error es nula la pérdida es nula, y cuando la entrada y la salida son estadísticamente independientes, la perdida es total, es decir, igual a la información a la entrada del canal.

Capacidad de canal para un canal digital

- Definición formal para un DMC

$$C = \max_{p_X(x_i)} I(X,Y)$$

Sus unidades son bits (o bits por uso de canal)

  - La maximización es respecto de las probabilidades de los símbolos de entrada
    - Hay que buscar las $p_X(x_i)$ que maximizan $I(X,Y)$

- Valores límite
  - $C \geq 0$, ya que $I(X,Y) \geq 0$
  - $C \leq \log M_X$, ya que $C = \max I(X,Y) \leq \max H(X) = \log M_X$
  - $C \leq \log M_Y$, por la misma razón

Ejemplo - Canal binario simétrico

- Modelo para canal digital binario con $BER = \varepsilon$
- Cálculo de la información mutua entrada / salida

$$\begin{aligned}
I(X,Y) =& H(Y) - H(Y|X) = H(Y) - \sum_{i=0}^{1} p_X(x_i) \cdot H(Y|X=x_i) \\
=& H(Y) - \sum_{i=0}^{1} p_X(x_i) \cdot \left[-\sum_{j=0}^{1} p_{Y|X}(y_j|x_i) \cdot \log p_{Y|X}(y_j|x_i) \right] \\
=& H(Y) - \sum_{i=0}^{1} p_X(x_i) \cdot [-\varepsilon \cdot \log(\varepsilon) - (1-\varepsilon) \cdot \log(1-\varepsilon)] \\
=& H(Y) - \sum_{i=0}^{1} p_X(x_i) \cdot H_b(\varepsilon) = H(Y) - H_b(\varepsilon)
\end{aligned}$$

- Cálculo de la capacidad de canal
  - Se busca el máximo de la información mutua
    - Para este canal, se obtiene cuando $H(Y)$ es máxima
    - $H(Y)$ es máxima cuando los símbolos de salida son equiprobables
    - Para este canal, ocurre cuando los símbolos de entrada son equiprobables

$$\boxed{C = 1 - H_b(\varepsilon)}$$

  - Probabilidades de X que maximizan $I(X,Y)$: $p_X(x_0) = p_X(x_1) = \frac{1}{2}$

Capacidad: problema de maximización con restricciones

- El problema del cálculo de la capacidad de un canal digital se puede plantear como un problema de maximización de una función con restricciones
  - Función a maximizar
    - Información mutua entre entrada y salida del canal $I(X, Y)$
  - Variables sobre las que se maximiza
    - Probabilidades de los símbolos de entrada $\{p_X(x_i)\}_{i=0}^{M_X-1}$
  - Restricciones
    - Valor de cada probabilidad

$$0 \leq p_X(x_i) \leq 1, \text{ para } i \in \{0, 1, \cdots, M_X - 1\}$$

    - Suma de todas las probabilidades

$$\sum_{i=0}^{M_X-1} p_X(x_i) = 1$$

- En general encontrar una solución analítica puede ser difícil
  - Soluciones analíticas sólo para canales "*sencillos*"
  - Cálculo mediante métodos numéricos utilizando ordenadores

Límites para la transmisión en un canal digital

- Un canal digital tiene una capacidad C bits/uso
  - Si se utilizan codigos de canal, los códigos prácticos (aquellos que permiten reducir la probabilidad de error de forma arbitraria) deberán tener una tasa de codificación menor que dicha capacidad

$$R < C$$

- Limitación práctica en términos de velocidad de transmisión efectiva cuando se utiliza codificación para protección frente a errores
  - Sistema diseñado para transmitir a R_b bits/s

$$R_b^{efectiva} = R \times R_b \text{ bits de infomación/s}$$

    - Límite en la tasa de transmisión efectiva

$$R_b^{efectiva} < C \times R_b \text{ bits de infomación/s}$$

Capacidad de canal para el canal gausiano

- Modelo de relación entrada salida en un canal gausiano

$$Y = X + Z$$

Z es una variable aleatoria gausiana, de media nula y varianza P_Z

- Capacidad de canal en las siguientes condiciones:
  - Potencia transmitida: P_X watt.
  - Ancho de banda: B Hz
    - Potencia de ruido: $P_Z = N_0 B$ watt.
- Cálculo a través de la información mutua

$$C = \max_{f_X(x) \mid E[X^2] \leq P_X} I(X, Y)$$

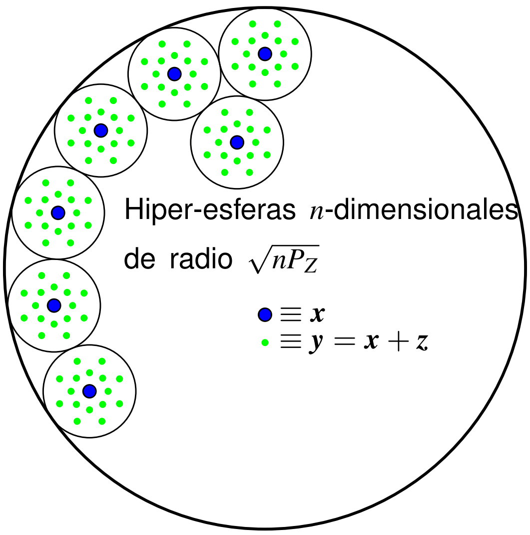
Restricción $E[X^2] \leq P_X$ dada por la limitación en potencia

- Resultado

$$C = B \cdot \log \left(1 + \frac{P_X}{N_0 B}\right) \text{ bits/s}$$

Se obtiene para $f_X(x)$ gausiana

Capacidad de canal para canal gausiano (II)



Hiper-esfera n-dimensional de radio $\sqrt{n(P_X + P_Z)}$

Capacidad sobre canal gausiano en las siguientes condiciones:

- Potencia transmitida: P_X watt.
- Ancho de banda: B Hz
  - Potencia de ruido: $P_Z = N_0 B$ watt.

Capacidad obtenida a partir del número de secuencias sin solapamiento

$$M_{ss} = \left(1 + \frac{P_X}{P_Z}\right)^{n/2}$$

$$C = \frac{\log M_{ss}}{n} = \frac{1}{2} \log \left(1 + \frac{P_X}{P_Z}\right)$$

$$C = \frac{1}{2} \cdot \log \left(1 + \frac{P_X}{N_0 B}\right) \text{ bits/uso}$$

$$\boxed{C = B \cdot \log \left(1 + \frac{P_X}{N_0 B}\right) \text{ bits/s}}$$

Evolución de la capacidad de canal

- Capacidad del canal gausiano

$$C = B \cdot \log\left(1 + \frac{P_X}{N_0 B}\right) \text{ bits/s}$$

- Depende de dos parámetros
  - Potencia de la señal transmitida, P_X
  - Ancho de banda disponible en Hz, B
- Capacidad de canal en función de la potencia transmitida P_X

$$\lim_{P_X \to \infty} C = \infty$$

  - Se puede aumentar C de forma arbitraria aumentando P_X
  - Aumento lineal de C requiere aumento exponencial de P_X
- Capacidad de canal en función del ancho de banda (B Hz)

$$\lim_{B \to \infty} C = \frac{P}{N_0} \log_2(e) = 1{,}44 \cdot \frac{P}{N_0}$$

  - El incremento de B no permite un incremento arbitrario de C

Límites para la transmisión en un canal gausiano

- Algunas definiciones
  - Tasa de transmisión binaria: R_b bits/s
  - Relación señal a ruido: $SNR = \frac{P_X}{P_Z} = \frac{P_X}{N_0 B}$
  - Tasa binaria (eficiencia) espectral: $\eta = \frac{R_b}{B}$ bits/s/Hz
  - Energía media por bit: $E_b = \frac{P_X}{R_b}$
  - Relación E_b/N_0: $\frac{E_b}{N_0} = \frac{P_X}{R_b N_0} = \frac{P_X}{N_0 B} \times \frac{B}{R_b} = \frac{SNR}{\eta}$
- Sistema de comunicaciones práctico

$$R_b < C \to R_b < B \cdot \log\left(1 + SNR\right) \text{ bits/s}$$

  - Dividiendo por B en ambos lados y reordenando

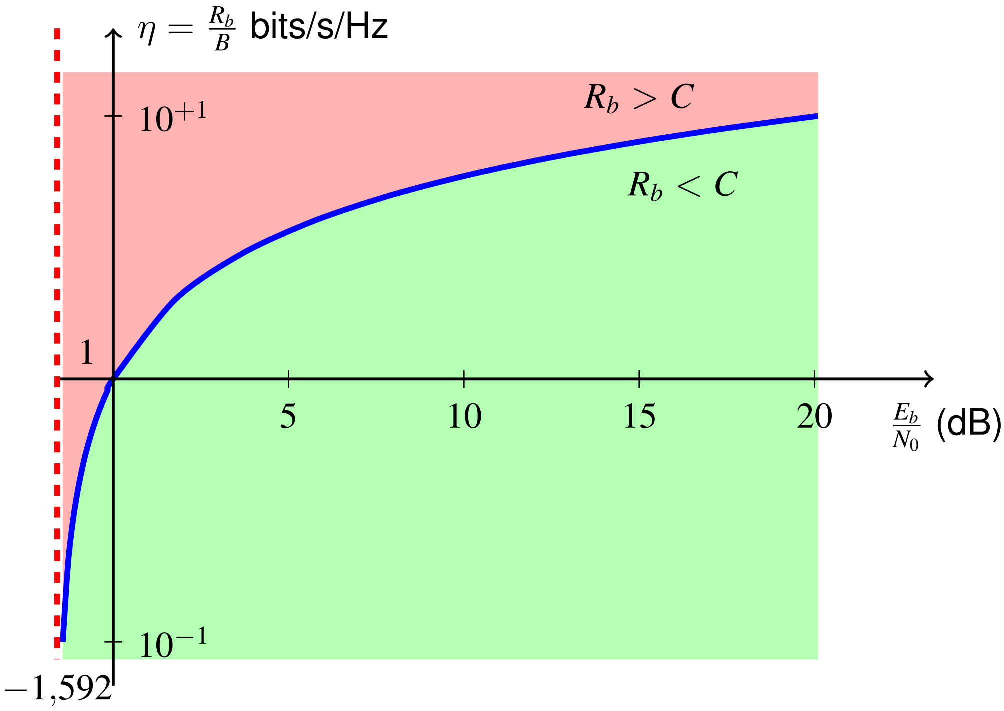
$$\eta < \log\left(1 + SNR\right), \quad \eta < \log\left(1 + \eta \cdot \frac{E_b}{N_0}\right)$$

$$SNR > 2^\eta - 1, \quad \frac{E_b}{N_0} > \frac{2^\eta - 1}{\eta}$$

Cuando $\eta \to 0$ $\quad \frac{E_b}{N_0} = \ln 2 = 0{,}693 \approx -1{,}6$ dB

Tasa binaria espectral frente a E_b/N_0

- Se representa sobre el plano η vs $\frac{E_b}{N_0}$ la curva $\frac{E_b}{N_0} = \frac{2^\eta - 1}{\eta}$
  - Divide el plano en dos regiones: sistemas con $R_b > C$ (prácticos) y con $R_b > C$



Relación señal a ruido normalizada

- Cota inferior para SNR

$$SNR > 2^\eta - 1$$

- Definición de SNR normalizada

$$SNR_{norm} = \frac{SNR}{2^\eta - 1}$$

- Cota inferior sobre SNR_{norm}

$$SNR_{norm} > 1 \text{ (0 dB)}$$