② 3º TEL. SUP. 0.30€

Comunicaciones Analógicas

Ingeniería de Comunicaciones. Créditos: 4,5 Obligatoria

OBJETIVOS FUNDAMENTALES

Presenta los fundamentos de la Teoría de la Comunicación. Introduce los conceptos y resultados principales de la Teoría de la Información. Describe y analiza las diferentes técnicas de modulación analógica, y destaca sus diferencias y similitudes. El concepto de equivalente paso bajo de señales paso banda es introducido y empleado como herramienta básica para su análisis

Asignaturas que se recomienda cursar con anterioridad: **Señales y Sistemas I y II.**

PROGRAMA

1. Teoría de la Información. Concepto de entropía. Fuente discreta sin memoria. Codificación de fuente: Primer teorema de Shannon. Códigos Huffman y Lempel-Ziv. Canales discretos sin memoria. Información mutua. Capacidad de un canal. Decisión óptima y ML. Segundo teorema de Shannon. Entropía diferencial de fuentes continuas. Canal Gaussiano. Ley de Shannon-Hartely.

2. Conceptos básicos de Teoría de la Señal. Señales y sistemas complejos. Propiedades de simetría de la TF. Autocorrelación y densidad espectral de potencia de señales y procesos complejos. Procesos cicloestacionarios. Señales y procesos paso banda: Señal analítica y equivalente paso bajo. Transformada de Hilbert. Procesos paso banda estacionarios.

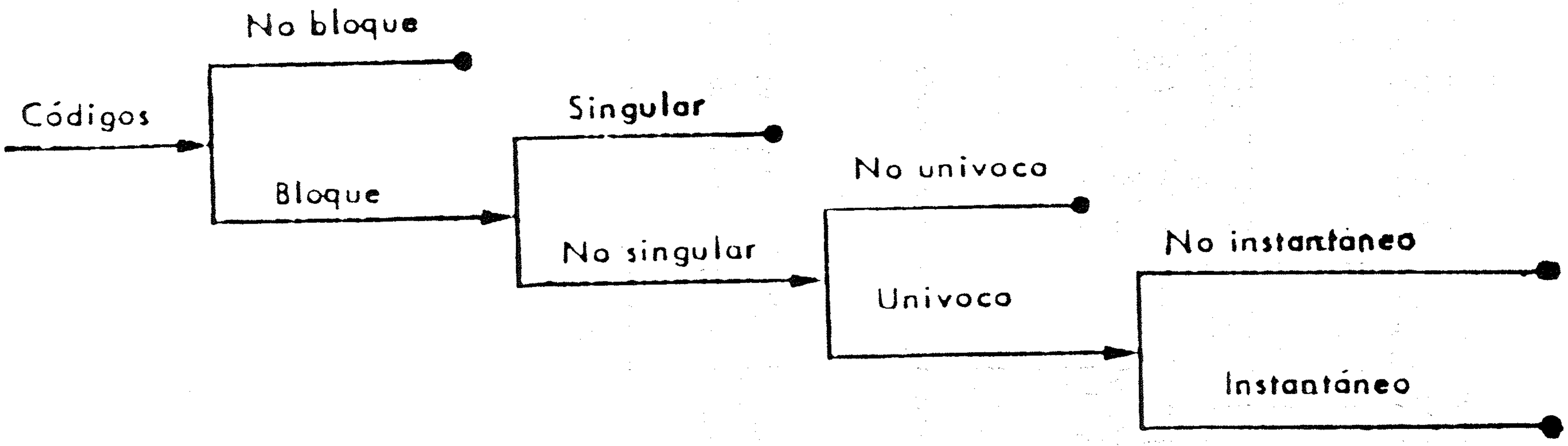
3. Modulaciones analógicas lineales y angulares. Modulaciones lineales: Doble Banda Lateral, Modulación de Amplitud, Banda Lateral Única, Banda Lateral Residual, Modulación en Cuadratura. Mezclador. Multiplexión en frecuencia. Modulaciones Angulares: Modulaciones de fase (PM) y frecuencia (FM). Espectro de la señal FM. Multiplicador de frecuencia. Demodulación de señales FM: Discriminador. Difusión de Radio y Televisión.

4. Efecto del ruido en el proceso de demodulación. Modelo de referencia. Calidad de un receptor banda base. Modulaciones lineales: Análisis de la demodulación coherente. Efecto de errores de fase y frecuencia en la demodulación. PLL. Análisis del detector de envolvente de AM. Análisis del Demodulador de FM. Efecto umbral. Receptor FM con realimentación (FMFB). Preénfasis/Deénfasis.

5. Modulaciones Analógica de pulsos. Codificación de fuentes analógicas. Modulación de Pulsos en Amplitud (PAM), Duración (PDM) y Posición (PPM). Multiplexión temporal. Análisis en ruido. Teoría de la Velocidad-Distorsión. Cuantificación escalar uniforme y logarítmica. Modulación por Pulsos Codificados (PCM).

BIBLIOGRAFÍA

J. G. PROAKIS, M. SALEHI: *Communications Systems Engineering*. Prentice-Hall. 1994.
parecidos SIMON HAYKIN. *Communication Systems*. Ed. John Wiley & Sons. 3º Edición, 1994.
importante → NORMAN ABRAMSON: *Teoria de la Información y Codificación*. Ed. Paraninfo.
A. PAPOULIS: *Probability, Random Variables and Stochastic Processes*. McGraw Hill. 3º Edición 1991



CÓDIGOS CLASIFICACIÓN

TEORÍA DE LA SEÑAL

$\mathrm{Sim}(x(t)) = x_e(t) = \frac{x(t)+x^*(-t)}{2}$ $\mathrm{Antisim}(x(t)) = x_o(t) = \frac{x(t)-x^*(-t)}{2}$

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi f t}\, dt$$

$x(-t) \xleftrightarrow{\mathcal{F}} X(-f)$ $x_R(t) \longleftrightarrow X_e(f)$ $x_e(t) \longleftrightarrow X_R(f)$

$x^*(-t) \longleftrightarrow X^*(f)$ $jx_I(t) \longleftrightarrow X_o(f)$ $x_o(t) \longleftrightarrow jX_I(f)$

$x^*(t) \longleftrightarrow X^*(-f)$

Estacionariedad: $m_x(t) = m_x$ $R_x(t, t-\tau) = R_x(\tau)$ $R_{xx^*}(t, t-\tau) = R_{xx^*}(\tau)$

Cicloestacionariedad: $m_x(t) = m_x(t+kT)$ $R_x(t,\tau) = R_x(t+kT,\tau)$ $R_{xx^*}(t,\tau) = R_{xx^*}(t+kT,\tau)$

↳ $\bar{m}_x = \frac{1}{T}\int_0^T m_x(t)\,dt$ $\bar{R}_x(\tau) = \frac{1}{T}\int_0^T R_x(t,\tau)\,dt$

$P_x(t) = E[|x(t)|^2] = R_x(t,0)$ $P_x = \langle E[|x(t)|^2] \rangle = E[\langle |x(t)|^2 \rangle]$

$\bar{R}_{xy}(\tau) = \bar{R}^*_{yx}(-\tau)$ $\bar{R}_x(\tau) = \bar{R}^*_x(-\tau)$ $\bar{R}_x(0) = \bar{P}_x$

$S_x(f) = \mathcal{F}\{R_x(\tau)\}$

$z(t) = x(t) + y(t) \Rightarrow \bar{R}_z(\tau) = \bar{R}_x(\tau) + \bar{R}_y(\tau) + \bar{R}_{xy}(\tau) + \bar{R}_{yx}(\tau) = \bar{R}_x(\tau) + \bar{R}_y(\tau) + 2\mathrm{Sim}(\bar{R}_{xy}(\tau))$

↳ $S_z(f) = S_x(f) + S_y(f) + 2\,\mathrm{Re}(S_{xy}(f))$

$\hat{x}(t) = x(t) * h_{TH}(t)$ $H_{TH}(f) = -j\,\mathrm{sign}(f)$ $h_{TH}(t) = \frac{1}{\pi t}$ $\hat{\hat{x}}(t) = -x(t)$

$h_{SPL}(t) = \delta(t) + j\frac{1}{\pi t}$ $H_{SPL}(f) = 2\,u(f)$ $R_{\hat{x}}(\tau) = R_x(\tau)$ $R_{x\hat{x}}(\tau) = -\hat{R}_x(\tau)$

$X_a(f) = 2X^+(f) = X_{eq}(f-f_0)$ $x_a(t) = x(t) * h_{SPL}(t) = x(t) + j\hat{x}(t) = x_{eq}(t)\,e^{j2\pi f_0 t}$

$X_{eq}(f) = X_a(f+f_0)$ $x_{eq}(t) = x_a(t)\,e^{-j2\pi f_0 t}$

$x_{eq}(t) = x_f(t) + jx_c(t)$

$-\hat{R}_x(\tau)$

$R_{x_a}(\tau) = 2R_x(\tau) - j2\,\mathrm{Impar}(-\hat{R}_x(\tau))$

$S_{x_a}(f) = 4S_x^+(f) = S_{x_{eq}}(f-f_0)$

$R_{x_a x_a^*}(\tau) = 0$

MODULACIONES LINEALES

DBL: $x_F(t) = z(t)$, $x_c(t) = 0$, $P_x = \frac{A^2}{2} P_z$, $X(f) = \frac{A}{2}\left(Z(f-f_c) + Z(f+f_c)\right)$

AM: $x_F(t) = 1 + m\, z(t)$, $x_c(t) = 0$, $P_x = \frac{A^2}{2}(1 + m^2 P_z)$, $X(f) = \frac{A}{2}\left(\delta(f-f_c) + \delta(f+f_c) + m Z(f-f_c) + m Z(f+f_c)\right)$

BLU: $x_F(t) = z(t)$, $x_c(t) = \pm \hat{z}(t)$, $P_x = A^2 P_z$, $X(f) = A\left(Z_+(f-f_c) + Z_-(f+f_c)\right)$

MODULACIONES ANGULARES

$x(t) = A\cos\left(2\pi f_c t + \phi(t)\right)$ $\xrightarrow{PM}$ $\phi(t) = \phi_\Delta z(t)$, $0 < \phi_\Delta \leq \pi$, $f(t) = f_c + \frac{1}{2\pi}\frac{d\phi(t)}{dt}$

$\xrightarrow{FM}$ $\phi(t) = 2\pi f_\Delta \int_0^t z(\tau)d\tau$, $f(t) = f_c + f_\Delta z(t)$

$W = 2(B + f_\Delta)$

EFECTO DEL RUIDO

DBL: $(S/N)_E = \frac{P_R}{2\eta B}$, $(S/N)_S = \frac{P_R}{\eta B}$

AM: $(S/N)_E = \frac{P_R}{2\eta B}$, $(S/N)_S = \frac{m^2 P_z}{1 + m^2 P_z}\frac{P_R}{\eta B}$

BLU: $(S/N)_E = \frac{P_R}{\eta B}$, $(S/N)_S = \frac{P_R}{\eta B}$

FM: $(S/N)_E = \frac{P_R}{\eta\, 2(D+1) B}$, $(S/N)_S = \frac{P_R}{\eta B}\, 3\underbrace{\left(\frac{f_\Delta}{B}\right)}_{D}^2 P_z$

$H_{PE}(f)\, H_{D-E}(f) = 1$

$\Downarrow$

$$\frac{(S/N)_{D-E}}{(S/N)_S} = \frac{1}{3}\frac{B^3}{\int_0^B |H_{D-E}(f)|^2 f^2\, df}$$

Formulario COMUNICACIONES ANALÓGICAS 3º I. TELECOMUNICACIÓN

$$I(A;B)=H(A)-H(A/B)=H(B)-H(B/A) \qquad H(A/B)=\sum_i^q\sum_j^r p(a_i,b_j)\log\frac{1}{p(a_i/b_j)}$$

$$C_{Uniforme}=\begin{Bmatrix} Max\ H(B) \\ \{p(a_i)\} \end{Bmatrix}+\sum_j p(b_j/a_i)\log p(b_j/a_i) \qquad C_{DébilSimétrico}=\log r+\sum_j p(b_j/a_i)\log p(b_j/a_i)$$

$$C=\frac{1}{2}\log\left(1+\frac{S}{N}\right) \qquad h_{gauss}(X)=\frac{1}{2}\log\left(2\pi e\sigma^2\right) \qquad C=\frac{1}{2}\int_{-W}^{W}\log\left(1+\frac{S_x(f)}{S_v(f)}\right)df$$

$\mathbf{R}_x(\tau)=\mathbf{R}_{x_R}(\tau)+\mathbf{R}_{x_I}(\tau)-2j\ \text{Impar}\left[\mathbf{R}_{x_Rx_I}(\tau)\right]$	$\mathbf{S}_x(f)=\mathbf{S}_{x_R}(f)+\mathbf{S}_{x_I}(f)+2\,\text{Im}\left[\mathbf{S}_{x_Rx_I}(f)\right]$
$\mathbf{R}_{xx^*}(\tau)=\mathbf{R}_{x_R}(\tau)-\mathbf{R}_{x_I}(\tau)+2j\ \text{Par}\left[\mathbf{R}_{x_Rx_I}(\tau)\right]$	$\mathbf{S}_{xx^*}(f)=\mathbf{S}_{x_R}(f)-\mathbf{S}_{x_I}(f)+2j\ \text{Re}\left[\mathbf{S}_{x_Rx_I}(f)\right]$

T.Hilbert: $h(t)=\dfrac{1}{\pi t}$; $H(f)=-j\ \text{sig}(f)$ $\qquad x_a(t)=x(t)+j\hat{x}(t)$; $x_{eq}(t)=x_a(t)\,e^{-j2\pi f_0 t}=x_F(t)+jx_C(t)$

Las relaciones generales entre las autocorrelaciones y las densidades espectrales son:

$\overline{R}_{x_a}(\tau)=2\left[\overline{R}_x(\tau)+j\hat{\overline{R}}_x(\tau)\right]$	$S_{x_a}(f)=4S_x^+(f)$
$\overline{R}_{x_{eq}}(\tau)=\overline{R}_{x_a}(\tau)e^{-j2\pi f_0 t}$	$S_{x_{eq}}(f)=S_{x_a}(f+f_0)$
$\overline{R}_x(\tau)=\dfrac{\text{Re}\left[\overline{R}_{x_a}(\tau)\right]}{2}$	$S_x(f)=\dfrac{S_{x_{eq}}(f-f_0)+S_{x_{eq}}(-f-f_0)}{4}$

Si x(t) es una señal aleatoria estacionaria, $x_F(t)$ y $x_C(t)$ son conjuntamente estacionarias y:

$R_{x_F}(\tau)=R_{x_C}(\tau)=\frac{1}{2}\text{Re}\left[R_{x_{eq}}(\tau)\right]$	$S_{x_F}(f)=S_{x_C}(f)=\frac{1}{2}\text{Par}\left[S_{x_{eq}}(f)\right]=S_x^+(f+f_0)+S_x^-(f-f_0)$
$R_{x_Fx_C}(\tau)=-\frac{1}{2}\text{Im}\left[R_{x_{eq}}(\tau)\right]$	$S_{x_Fx_C}(f)=\frac{j}{2}\text{ImPar}\left[S_{x_{eq}}(f)\right]=j\left[S_x^+(f+f_0)-S_x^-(f-f_0)\right]$

Si $x_F(t)$ y $x_C(t)$ son señales aleatorias conjuntamente estacionarias, x(t) es en general cicloestacionaria:

$R_x(t,\tau)=\text{Re}\left[\dfrac{R_{x_a}(\tau)+R_{x_ax_a^*}(t,\tau)}{2}\right]$	$R_{x_a}(\tau)=R_{x_{eq}}(\tau)e^{j2\pi f_0\tau}$ $R_{x_ax_a^*}(t,\tau)=R_{x_{eq}x_{eq}^*}(\tau)e^{-j2\pi f_0\tau}e^{j4\pi f_0 t}$

	(S/N)E	(S/N)S	GP
AM	$\delta/2$	$\delta\left(\dfrac{m^2P_z}{1+m^2P_z}\right)$	$2\left(\dfrac{m^2P_z}{1+m^2P_z}\right)$
DBL	$\delta/2$	δ	2
BLU	δ	δ	1
MQ	$\delta/2$	$\delta/2$	1
FM	$\dfrac{\delta}{2(D+1)}$	$3.D^2P_z\delta$	$6D^2(D+1)P_z$

$$\delta=\frac{P_R}{\eta B} \qquad D=\frac{f_\Delta}{B}$$

$$\phi_{PM}(t)=\Phi_\Delta z(t)$$

$$f_{PM}(t)=f_C+\frac{1}{2\pi}\phi_\Delta\frac{dz(t)}{dt}$$

$$\phi_{FM}(t)=2\pi f_\Delta\int_0^t z(\tau)d\tau$$

$$f_{FM}(t)=f_C+f_\Delta z(t)$$

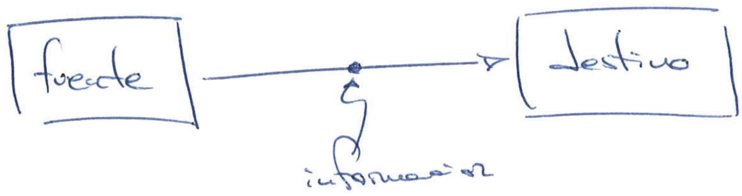
Notas: 11 Julio

Rev: 13/7 12:00 | 14/7 16:00

Tema 1: Teoría de la Información

1.- Introducción

· <u>Comunicación</u>: intercambio de información (mensaje) entre dos entes "inteligentes"



· <u>Telecomunicación</u>: comunicación entre dos puntos alejados

La <u>Teoría de la Información</u> es aplicable también a la grabación y posterior reproducción de la información.

Podemos decir que esta información se concreta en forma de <u>mensajes</u>, compuestos por una secuencia de elementos de un alfabeto.

· <u>Mensaje</u>: concatenación de una serie de símbolos que pertenecen a un conjunto finito

mensaje: $m(n) \in A$

eje discreto

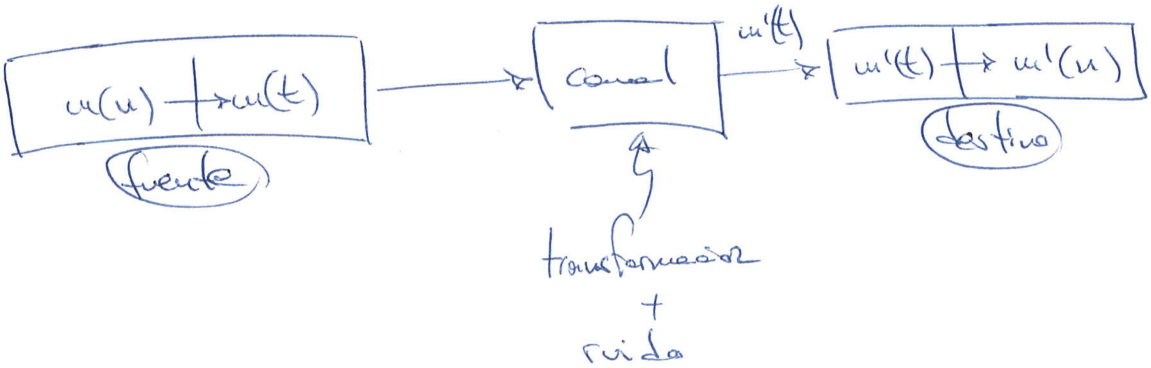
señal

Cuando alguien quiere comunicar, se necesita un soporte físico para el mensaje:

- **Materialización de la información:**
  - variación de ondas de presión acústica (ej: voz)
  - variación de corriente/tensión en un conductor (ej: telégrafo)
  - variación de magnitud de campo electromagnético (ej: radio)

  $m(t)$ continua

La fuente se puede ver como un generador de mensajes:



Dos tipos de fuentes:
- fuente digital: genera señal digital $m(n)$
- fuente analógica: genera señal continua $m(t)$

· Sistema de comunicaciones genérico:

$$\boxed{\text{fuente}} \xrightarrow[m(n)]{m(t)} \boxed{Tx} \xrightarrow[x(t)]{y(t)} \boxed{\text{canal}} \xrightarrow{y'(t)} \boxed{Rx} \xrightarrow[m'(n)]{m'(t)} \boxed{\text{destino}}$$

(digital o analógica) (analógica) (analógica) (digital o analógica)

El Tx acomoda la señal de la fuente (analógica o digital) al canal por el que se va a transmitir. Genera perturbaciones de una magnitud física susceptible de ser propagada en la distancia. Esto es, produce una señal continua analógica, tanto al transmitir una señal digital como analógica.

El canal realiza una transformación de la señal que le introduce el Tx, y añade una señal espúrea, concretamente ruido, $n(t)$.

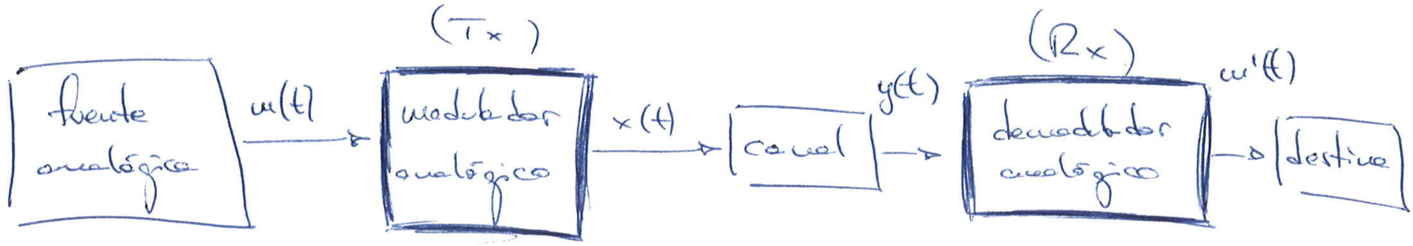
El dispositivo receptor tiene como objetivo tratar de recuperar el mensaje original de la fuente.

Los criterios de calidad difieren según el sistema de comunicación sea analógico o digital:

- digital: probabilidad de error $\rightarrow P_e = P(m(n) \neq m'(n))$
- analógico: relación señal-ruido

$$\rightarrow S/N = \frac{P(m(t))}{P(m(t) - m'(t))} \quad [dB]$$

· Sistema de comunicaciones analógicas:



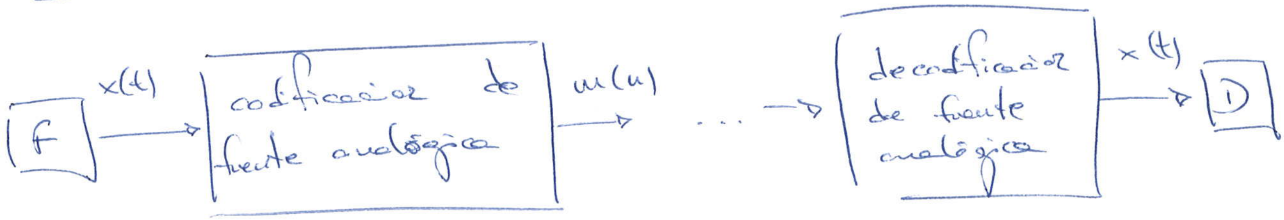
modulaciones
- lineales (AM, DBL, ...): ancho de banda ↓, S/N ↓
- angulares (FM, PM): ancho de banda ↑, S/N ↑

Elección de modulaciones:

- ancho de banda disponible
- tipos de ruido (sensibilidad, etc.)
- dispositivos amplificadores

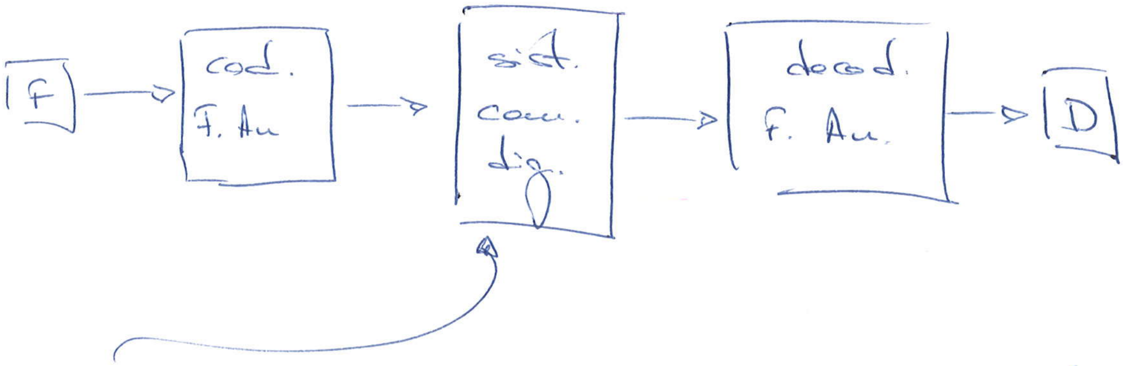
⇒ compromiso entre ancho de banda y S/N

· Codificación de fuentes analógicas:



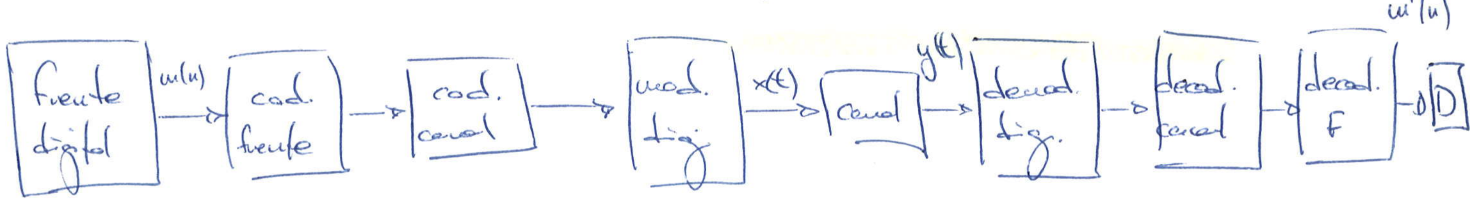
Desde el punto de vista matemático, la pérdida de información es infinita, pero desde el punto de vista del receptor, es pequeña

La codificación es útil porque el destinatario es incapaz de percibir toda la información de la señal que produce la fuente. Además, aunque el destinatario perciba diferencias, no le importa. Ej: voz



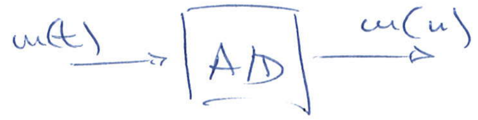
con un sistema de comunicación digital se puede controlar mucho mejor la calidad del sistema global.

· <u>Sistema de comunicaciones digitales:</u>



· <u>Cantidad de información</u> que proporciona una fuente:

- número mínimo de símbolos necesarios para representar la información (ej: bits) → <u>señales digitales</u>
- información continua/infinita ⇒ codificación A/D → <u>señales analógicas</u>



se pierde algo de información, pero puede que no afecte si el destino tiene resolución limitada.

· Teoría de la Información:

1er Teorema de Shannon: manera más compacta de transmitir información, o de usar el mínimo número de símbolos para representarla

2º Teorema de Shannon: dadas las características de un canal, conocer la máxima información que se puede transmitir por el mismo, sin errores.

El padre de la Teoría de la Información fue Shannon, desarrollada a finales de los 40 principios de los 50.

2.- CONCEPTO DE ENTROPÍA

· entropía: forma de medir la información

procesos estocásticos

Un mensaje es un suceso con cierta probabilidad de aparición

La entropía se utiliza para dar una medida objetiva, asociada a la incertidumbre (probabilidad)

Se tiene un experimento con una serie de resultados con sus probabilidades de ocurrencia:

$$S' = \{ s_1, s_2, \ldots, s_N \}$$

$$p_1 \quad p_2 \quad \ldots \quad p_N \qquad \sum_{i=1}^{N} p_i = 1$$

Cuando se produce un resultado, se puede obtener la información que da ese resultado, o la incertidumbre que disminuye.

<u>Requisitos</u> de la medida de información:

a) $I(s_i)$ = cantidad de información del resultado s_i

b) $p_i > p_j \Rightarrow I(s_i) < I(s_j)$

Cuanto mayor es la probabilidad de un suceso, menor incertidumbre tiene, por lo que su cantidad de información es menor.

c) $p_i = 1 \Rightarrow I(s_i) = 0$

d) Si se hacen 2 experimentos o repeticiones independientes, s_i y s_j:

$$S \times S = \{s_i s_j\} \quad i,j = 1:N \Rightarrow N^2 \text{ posibles resultados}$$

$$p(s_i s_j) = p_i \cdot p_j \Rightarrow I(s_i s_j) = I(s_i) + I(s_j)$$

e) $I(s_i) \geq 0$

Una medida que cumple estos requisitos es:

$$\boxed{I(s_i) = \log \frac{1}{p_i}} = -\log p_i$$

- $I(s_i)$ es inversamente proporcional a p_i
- media de la información: es más interesante que la información de un solo experimento

$$E[I(s_i)] = \sum_{i=1}^{N} p_i I(s_i) = \sum_{i=1}^{N} p_i \log 1/p_i = -\sum_{i=1}^{N} p_i \log p_i$$

Esto es lo que se conoce como ENTROPÍA:

$$\boxed{H(S') = E[I(s_i)] = -\sum_{i=1}^{N} p_i \log p_i}$$

La entropía nos dice la cantidad de incertidumbre que se elimina cuando se produce un experimento.

- Unidades:

(propiedad básica de los logaritmos: $\log_a x = \frac{\log_b x}{\log_b a}$)

a) $\log_2 \longrightarrow$ bits

b) $\ln \longrightarrow$ nats

c) $\log_{10} \longrightarrow$ Hartleys

1 Hart = 3,32 bits

1 nat = 1,44 bits

$$\boxed{\text{Por defecto, } \log = \log_2}$$

- Propiedad importante de la entropía:

$$\boxed{0 \leq H(S') \leq \log N}$$

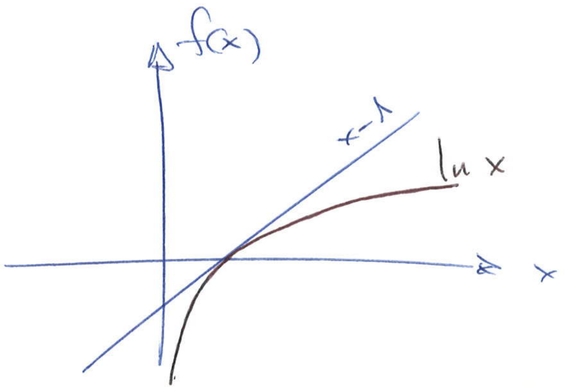
$H(S') = \log N \iff p_i = \frac{1}{N}$

La entropía es también una medida de la aleatoriedad del experimento.

Se relaciona por tanto con la entropía en termodinámica, que corresponde con el desorden en los materiales.

- Demostración de la propiedad:

$\Rightarrow \ln x \leq x - 1 \quad \forall x \geq 0$



Se tienen
- N símbolos con probabilidades p_i
- N símbolos con probabilidades $q_i \rightarrow \sum_{i=1}^{N} q_i = 1$

$$\sum_{i=1}^{N} p_i \log \frac{q_i}{p_i} = \frac{1}{\ln 2} \sum_{i=1}^{N} p_i \ln \frac{q_i}{p_i} \leq \frac{1}{\ln 2} \sum_{i=1}^{N} p_i \left(\frac{q_i}{p_i} - 1 \right)$$

$$= \frac{1}{\ln 2} \sum_{i=1}^{N} (q_i - p_i)$$

$$= \frac{1}{\ln 2} \left(\sum_{i=1}^{N} q_i - \sum_{i=1}^{N} p_i \right) = 0$$

Es decir, la desigualdad queda:

$$\sum_{i=1}^{N} p_i \log \frac{q_i}{p_i} = \sum_{i=1}^{N} p_i \log \frac{1}{p_i} - \sum_{i=1}^{N} p_i \log \frac{1}{q_i} \leq 0$$

$$\Rightarrow H(S) = \sum_{i=1}^{N} p_i \log \frac{1}{p_i} \leq \sum_{i=1}^{N} p_i \log \frac{1}{q_i}$$

Tomando $q_i = 1/N \Rightarrow \underline{H(S) \leq \log N}$

Como $\ln x = x - 1$ sólo en $x = 1$, entonces para que $H(S) = \log N$, debe ser $p_i = q_i = \frac{1}{N}$

3.- FUENTES DISCRETAS SIN MEMORIA

$$\boxed{F} \longrightarrow s(n) \qquad s(n) \in S$$

Una <u>fuente discreta sin memoria</u> es una entidad que genera de forma aleatoria una sucesión de símbolos de un alfabeto, a intervalos regulares de duración T (segundos).

Cada vez que se genera un símbolo, se produce un resultado del experimento.

<u>Sin memoria</u> $\Rightarrow$ las repeticiones son estadísticamente independientes

Se puede definir la entropía de la fuente asociada a la entropía del experimento S', $H(S')$: cantidad de información por cada símbolo que sale de la fuente

$$H(S') = \sum_{i=1}^{N} p_i \log 1/p_i \quad [\text{bits}]$$

Se puede dotar de temporalidad a este término, teniendo en cuenta que la fuente genera 1 símbolo cada T segundos:

· velocidad de la fuente:

$$\boxed{v(S') = \frac{H(S')}{T} \quad [\text{bps}]}$$

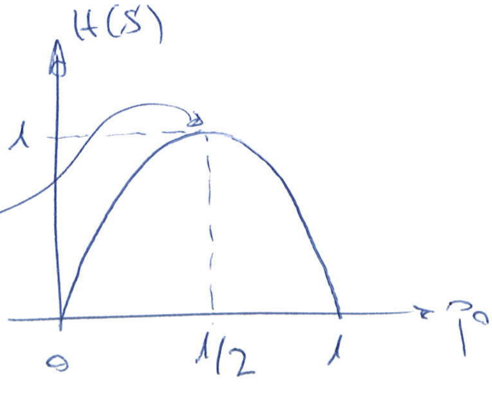
Ejemplo: fuente binaria sin memoria

$$S' = \{s_0, s_1\} \; ; \; P = \{p_0, p_1\}$$

$$p_0 + p_1 = 1$$

$$p_1 = 1 - p_0$$

$$H(S') = p_0 \log 1/p_0 + (1-p_0) \log \frac{1}{1-p_0}$$



máxima información cuando los símbolos son equiprobables: $p_0 = p_1 = 1/2$

Cuando la entropía depende de una sola variable (fuente binaria), se la puede llamar <u>función entropía</u>:

$$\boxed{H(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}}$$

<u>Ejemplo:</u> fuente de 4 símbolos → $S = \{0, 1, 2, 3\}$

$P = \{1/2, 1/4, 1/8, 1/8\}$

$T = 1\,ms$

<u>Entropía:</u> $\boxed{H(S) = \sum_{i=1}^{4} p_i \log \frac{1}{p_i} = \frac{7}{4} \text{ bits/símbolo}}$

<u>Velocidad:</u> $\boxed{V = \frac{H(S)}{T} = 1{,}75 \text{ kbps}}$

- Extensión de una fuente sin memoria:

$\boxed{F} \rightarrow s(n) \in S = \{s_i\}$

Consiste en agrupar los símbolos que produce la fuente.

Por ejemplo, una extensión de orden 2 agrupa los símbolos de 2 en 2:

$\underbrace{s_i s_j}_{s_{ij}} \underbrace{s_k s_l}_{s_{kl}}$ → producirá la mitad de símbolos

$$s_{ij} \in S \times S = \{s_i, s_j\}$$

Como es una fuente sin memoria:

$$\boxed{P(s_{ij}) = p_i p_j = p_{ij}}$$

· Entropía: $H(S^2) = H(S) \cdot 2$ ← los símbolos aportan el doble de información

$$\boxed{H(S^n) = n \cdot H(S)}$$

La velocidad permanece constante:

$$v = \frac{H(S^n)}{n \cdot T} = \frac{n \cdot H(S)}{n \cdot T} = \frac{H(S)}{T}$$

4.- CODIFICACIÓN

4.1.- CÓDIGOS BLOQUE

$$\boxed{F} \xrightarrow{s(u)} \boxed{\text{codificador}} \longrightarrow c(u) \qquad \begin{pmatrix} s_i(u) \in S' \\ c(u) \in C \end{pmatrix}$$

· código: correspondencia de los símbolos del alfabeto fuente, $s_i \in S'$, con una secuencia de símbolos del alfabeto código, $c_i \in C$, que tienen una longitud determinada, l_i

· objetivo de la codificación de fuente: obtener una representación compacta de la información

- Clasificación de los códigos: $(*1*)$

· código bloque: $\forall s_i \longrightarrow$ secuencia fija de símbolos del alfabeto código

Nº de símbolos que forman la secuencia = longitud de la palabra código (l_i)

· código no singular: un código bloque es no singular cuando todas las palabras código son distintas

$$i \neq j \Rightarrow c_i \neq c_j$$

- <u>código unívoco</u> (AKA "unívocamente decodificable): secuencias diferentes del alfabeto fuente dan lugar a secuencias diferentes del alfabeto código
  ⇒ codificación sin ruido, invertible

- <u>código instantáneo</u>: se puede decodificar una palabra código a medida que lleguen los símbolos correspondientes a cada palabra. No es necesario esperar a recibir el símbolo siguiente al final de una palabra código

- <u>código prefijo</u>: ninguna palabra código contiene a otra ⇒ una palabra no es prefijo de otra.
  Un código prefijo también es unívoco e instantáneo.

Un código de longitud fija(*) y no singular siempre es unívoco e instantáneo.

(*) ⇒ $l_i = l \ \forall i$

4.2.- Inecuación de Kraft - McMillan

Sean $l_1, l_2, \dots, l_q \in \mathbb{N}$ las longitudes de las palabras código de un código unívoco y r el número de símbolos del alfabeto código (ej: binario => r = 2)

Si se cumple

$$\boxed{\sum_{i=1}^{q} r^{-l_i} \leq 1}$$

entonces existe un código instantáneo cuyas palabras código tienen esas longitudes, $l_1, l_2, \dots, l_q$.

Todo código instantáneo debe cumplir esa expresión.

→ tanto los códigos unívocos como los instantáneos cumplen la inecuación, por lo que se puede (y es conveniente) hacer instantáneos los unívocos

Corolario: si se tiene un código unívoco con longitudes $l_1, l_2, \dots, l_q$, entonces existe un código instantáneo con esas mismas longitudes.

4.3.- LONGITUD MEDIA DE UN CÓDIGO

Se define como el número medio de símbolos del alfabeto código que son necesarios para representar un símbolo del alfabeto fuente.

$$\boxed{L = \sum_{i=1}^{q} p_i \cdot l_i}$$

Da una idea de lo bien (o mal) que se hace la codificación de fuente.

- Código compacto: código único de longitud media mínima

- Propiedad: $\boxed{H_r(S) \leq L}$

4.4.- PRIMER TEOREMA DE SHANNON

Partimos de la última expresión, e intentamos demostrarlo:

$$H_r(S) = \frac{H(S)}{\log r} \leq L$$

la entropía de la fuente en unidades r-arias es siempre menor o igual que la longitud media del código

$$H_r(S) = \sum_{i=1}^{q} p_i \log_r \frac{1}{p_i}$$

Como ya vimos, $\underline{\sum_{i=1}^{q} p_i \log_r \frac{1}{p_i} \leq \sum_{i=1}^{q} p_i \log_r \frac{1}{q_i}}$ $\quad 0 \leq q_i \leq 1$, $\sum_{i=1}^{q} q_i = 1$

Tomamos $q_i = \dfrac{r^{-l_i}}{\sum_{i=1}^{q} r^{-l_i}}$ con $\sum_{i=1}^{q} r^{-l_i} \leq 1$

$$\sum p_i \log_r 1/p_i \leq \sum p_i \log_r \frac{\sum_{i=1}^{q} r^{-l_i}}{r^{-l_i}} = \sum p_i \log_r \sum_{i=1}^{q} r^{-l_i} - \sum p_i \log_r r^{-l_i}$$

$$= \sum_{i=1}^{q} p_i \log_r \underbrace{\sum_{i=1}^{q} r^{-l_i}}_{\leq 1} + \sum_{i=1}^{q} p_i l_i$$

$\log_r \{\leq 1\} \leq 0$

$$\Rightarrow \underline{\sum_{i=1}^{q} p_i \log_r 1/p_i} \leq \sum_{i=1}^{q} p_i l_i + \{\leq 0\} = \underline{L + \{\leq 0\}}$$

$H_r(S) \leq L + \{\text{cti negativa}\} \Rightarrow \boxed{H_r(S) \leq L}$

¿Es alcanzable $H_r(S) = L$?

$$\rightarrow \begin{cases} q_i = p_i \\ \sum_{i=1}^{q} r^{-l_i} = 1 \end{cases} \Rightarrow p_i = r^{-l_i} \Rightarrow \boxed{l_i = -\log_r p_i}$$

distribución probabilística

en general no se cumple, ya que da un número real, y debe ser $l_i \in \mathbb{N}$

$$\boxed{H_r(S) = L \iff l_i = I(s_i) = -\log_r p_i}$$

Ejemplo: $S = \{s_1, s_2, s_3, s_4\}$

$P = \{1/2, 1/4, 1/8, 1/8\}$

Alfabeto código binario: $C = \{0, 1\}$

$l_1 = -\log p_1 = 1 \longrightarrow$ ej: 0

$l_2 = -\log p_2 = 2 \longrightarrow 10$

$l_3 = -\log p_3 = 3 \longrightarrow 110$

$l_4 = -\log p_4 = 3 \longrightarrow 111$

$$\begin{cases} H(S) = \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{2}{8} = \frac{7}{4} \text{ bits} \\ L = \sum_{i=1}^{4} p_i l_i = \frac{7}{4} \text{ bits} \end{cases}$$

Si conseguimos $l_i = -\log p_i \Rightarrow H(S) = L$, el código es compacto, de longitud media mínima

- Si $-\log p_i$ no es entero:

a) redondear: $l_i = \text{ceil}(-\log p_i)$

- El código no tendrá longitud media mínima
- Cumple $\sum r^{-l_i} \leq 1$
- También cumple $-\log_r p_i \leq l_i \leq -\log_r p_i + 1$

$$\downarrow$$

$$\begin{cases} -\log_r p_1 \leq l_1 \leq -\log_r p_1 + 1 \\ -\log_r p_2 \leq l_2 \leq -\log_r p_2 + 1 \\ \vdots \\ -\log_r p_q \leq l_q \leq -\log_r p_q + 1 \end{cases}$$

sumando y multiplicando por p_i

$$\downarrow$$

$$-\sum p_i \log_r p_i \leq \sum p_i l_i \leq -\sum p_i \log_r p_i + \overbrace{\sum p_i}^{1}$$

$$\Rightarrow \boxed{H_r(S) \leq L < H_r(S) + 1}$$

Para las extensiones se cumple igual:

$$\boxed{H_r(S^n) \leq L_n \leq H_r(S^n) + 1}$$

Como $H(S^n) = n \cdot H(S)$, al extender la fuente es posible aproximar la longitud media al mínimo teórico:

$$n \cdot H_r(S) \leq L_n \leq n \cdot H_r(S) + 1$$

dividir por n

$$\boxed{H_r(S) \leq \frac{L_n}{n} \leq H_r(S) + \frac{1}{n}}$$

$\frac{L_n}{n}$ es una medida del número medio de símbolos del alfabeto código utilizado para codificar un símbolo del alfabeto fuente considerando la fuente original, no la extensión

- Primer teorema de Shannon: $\boxed{\lim_{n \to \infty} \frac{L_n}{n} = H_r(S)}$

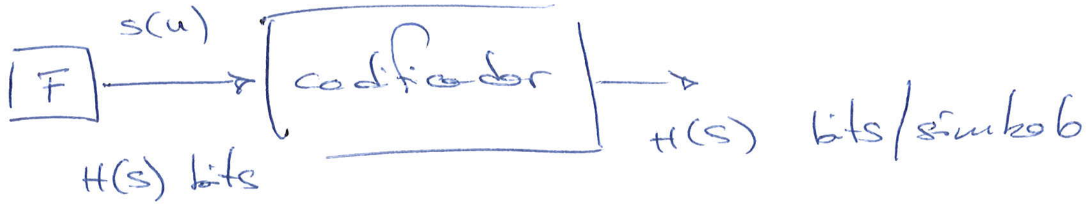
L_n = longitud media de la fuente extendida

$H_r(S)$ = entropía de la fuente sin extender

- Es una medida de lo bien que se puede codificar una fuente.

- Haciendo muchas extensiones se puede alcanzar el límite.

  → problema: enorme complejidad del código resultante



Si $H(S) = 100$ bits/símbolo y se generan 1000 símbolos, la cota máxima de extensión son 10^5 bits (100×1000)

4.5.- CÓDIGOS HUFFMAN

Son códigos compactos: para una fuente F, con alfabeto S, genera códigos unívocos de longitud mínima, para una extensión dada.

La codificación Huffman es un proceso de construcción de códigos compactos e instantáneos.

Para llevar a cabo el proceso, es necesario conocer los dos alfabetos y las probabilidades de ocurrencia de los símbolos de la fuente.

- Pasos: (*2*)

1.- Se colocan los símbolos ordenados por probabilidades decrecientemente

2.- Reducción de fuente hasta tener un número de símbolos igual al tamaño del alfabeto código

Para hacer la reducción, se agrupan tantos elementos como símbolos del alfabeto código, hasta tener ese mismo número de símbolos.

3.- Codificación: asignar símbolos del alfabeto código a lo que resulta, hacia atrás, añadiendo símbolos a los que vengan de agrupamientos y dejando como están los que no.

Se garantiza:

- que por la forma en que se codifica, el código es prefijo
- se añaden menos bits a los símbolos con menor probabilidad => longitud media mínima.

¡Ojo!: para un alfabeto código con r símbolos, se agrupan los elementos de r en r al hacer la reducción de fuente, y ésta termina cuando se tengan r símbolos. Las probabilidades se suman al agruparlos

Ejemplo: $S' = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9\}$

$C' = \{0, 1, 2\}$

S_i	P_i	
s_1	1/2	0
s_2	1/4	1
s_3	1/8	20
s_4	1/16	21
s_5	1/32	220
s_6	1/64	221
s_7	1/128	2220
s_8	1/256	2221
s_9	1/256	2222

$1/64 \quad 222$

$1/16 \quad 22$

$1/4 \quad 2$

$s_1 \rightarrow 0$ | $s_4 \rightarrow 21$ | $s_7 \rightarrow 2220$

$s_2 \rightarrow 1$ | $s_5 \rightarrow 220$ | $s_8 \rightarrow 2221$

$s_3 \rightarrow 20$ | $s_6 \rightarrow 221$ | $s_9 \rightarrow 2222$

Si no hay símbolos suficientes para agruparlos de r en r, se añaden al final símbolos "ficticios" con probabilidad nula. Esto suele ocurrir para $r \geq 3$

4.6.- RENDIMIENTO DE UN CÓDIGO

- Definición:

$$\boxed{\eta = \frac{H_r(S)}{L} \leq 1}$$

Al codificar extensiones de la fuente, $\eta \to 1$

Ejemplo: $S = \{s_0, s_1\}$ $\quad P = \{3/4, 1/4\}$

$C = \{0, 1\}$

$H(S) = 0{,}811$ bits $\qquad \Rightarrow \begin{cases} s_0 \to 0 \\ s_1 \to 1 \end{cases}$

$$L_1 = 1 \Rightarrow \eta_1 = 0{,}811$$

Extensión de orden 2:

$s_0 s_0 \to 9/16 \to 0$

$s_0 s_1 \to 3/16 \to 10$

$s_1 s_0 \to 3/16 \to 110$

$s_1 s_1 \to 1/16 \to 111$

$L_2 = \frac{27}{16}$

$\eta_2 = \frac{2H(S)}{L_2} = 0{,}961$

n	1	2	3	4
η	0,811	0,961	0,985	0,991

$$\boxed{F} \xrightarrow{1 \text{ bit}} \boxed{COD} \longrightarrow 0{,}811 \text{ bit}$$

Esta es la situación cuando se llega a $\eta = 1$

4.7.- ALGORITMO DE LEMPEL-ZIV

A veces no se conocen a priori las probabilidades de ocurrencia de los símbolos del alfabeto fuente, y otras veces la fuente tiene memoria.

Por ejemplo, una fuente que genera texto escrito

En una fuente con memoria $\Rightarrow H(S) \downarrow$

Lo ideal es tener codificadores ciegos, que funcionen bien con cualquier tipo de fuente

$\Rightarrow$ algoritmo (o códigos) de Lempel-Ziv

- No son códigos bloque
- Se pueden aplicar a cualquier fuente sin conocer su estadística
- Funcionan bien con fuentes con y sin memoria, mientras que Huffman sólo sin memoria
- Su rendimiento $\to 1$ asintóticamente, cualesquiera sean las estadísticas de la fuente
- Se usa en técnicas de compresión de ficheros (ZIP, ARJ,...)

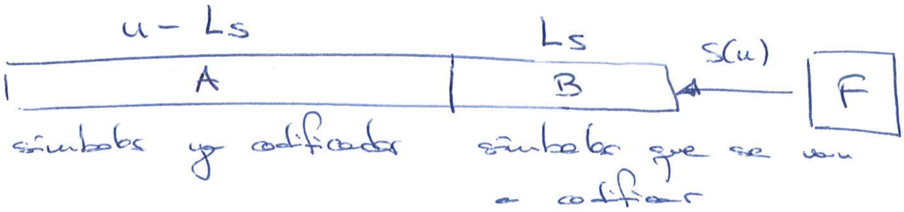
- Características:

1.- Codificador pensado para el caso en que el alfabeto código y el fuente son idénticos:

$C \equiv S$ (Normalmente binario)

2.- Codifica secuencias de símbolos del alfabeto fuente de longitud variable ($\leq L_s$) palabras código de longitud fija (L_c)

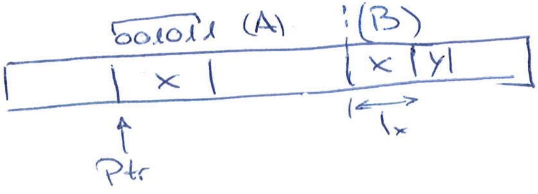
3.- Para el proceso de codificación se utiliza un registro de desplazamiento de u símbolos



$L_s \equiv$ longitud máxima de las palabras del alfabeto fuente

$$L_c = 1 + \mathrm{ceil}(\log(u - L_s)) + \mathrm{ceil}(\log(L_s))$$

• extensión reproducible: buscar en el registro B secuencias de símbolos que haya en alguna parte del registro A.



$X = 00101$

$Y = 0$

La secuencia de B se codifica entonces como un puntero a la posición de A en la que aparece, junto a la longitud de la secuencia encontrada y el bit siguiente (Y):

Ptr	lx	Y

- puntero: ceil (log (u - Ls)) bits $= L_{ptr}$
- longitud l_x: ceil (log (Ls)) bits $= L_{l_x}$
- bit siguiente: 1 bit

⇓

$$\boxed{L_c = 1 + \text{ceil}(\log(u - L_s)) + \text{ceil}(\log(L_s))}$$

Funciona bien con valores de u y Ls grandes

Ejemplo: $u - L_s = 4096$ bits

$L_s = 128$ bits

$L_{ptr} = 12$ bits

$L_{l_x} = 7$ bits

$\} L_c = 20$ bits

$u - L_s = 256$ bits

$L_s = 256$ bits

$L_{ptr} = 8$ bits

$L_{l_x} = 8$ bits

$\} L_c = 17$ bits

para codificar secuencias de hasta 256 bits

5.- CANALES DISCRETOS SIN MEMORIA

5.1.- INTRODUCCIÓN

- alfabeto de entrada: $A = \{a_1, a_2, \dots, a_q\}$
- alfabeto de salida: $B = \{b_1, b_2, \dots, b_r\}$
- probabilidades condicionales del canal: $p(b_j/a_i)$

  $\equiv$ probabilidad de recibir b_j si se ha recibido a_i

__modelado:__

$$a_i \longrightarrow \underset{\underline{\underline{P}}}{\overset{\text{(canal)}}{\boxed{p(b_j/a_i)}}} \longrightarrow b_j$$

$$\underline{\underline{P}} = \left[p(b_j/a_i) \right] = \begin{bmatrix} p_{11} & \cdots & p_{1r} \\ \vdots & & \vdots \\ p_{q1} & \cdots & p_{qr} \end{bmatrix}$$

filas: a_i

columnas: b_j

$$p_{ij} = p(b_j/a_i)$$

Son sin memoria porque la relación entre las a_i y las b_j es instantánea

__Propiedades:__

- la suma de las probabilidades de una fila será 1

$$\sum_{j=1}^{r} p(b_j/a_i) = 1 \quad \forall i$$

- dadas las probabilidades $p(a_i)$, se puede calcular $p(b_j)$:

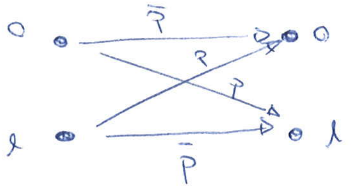
$$p(b_j) = \sum_{i=1}^{q} p(a_i)\, p(b_j/a_i)$$

- dadas las probabilidades $p(b_j)$, <u>no</u> se puede calcular $p(a_i)$

<u>ejemplo:</u> canal binario simétrico (BSC)

$$p(b_j/a_i) = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$ con p = probabilidad de error

$\bar{p} = 1 - p$



Si $p = 1/2 \Rightarrow p(b_j/a_i) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$

$p(a_i = 0) = w \Rightarrow p(a_i = 1) = \bar{w}$

$p(b_j = 0) = w\frac{1}{2} + \bar{w}\frac{1}{2} = \frac{1}{2}$

$p(b_j = 1) = w\frac{1}{2} + \bar{w}\frac{1}{2} = \frac{1}{2}$

la probabilidad de salida puede ser, como en este caso, independiente de las probabilidades de entrada

- dadas las probabilidades $p(a_i)$, podemos calcular $p(a_i/b_j)$ (probabilidad de que el emisor haya enviado a_i si se recibe un b_j): [<u>Teorema de Bayes</u>]

$$p(a_i/b_j) = \frac{p(a_i)\,p(b_j/a_i)}{\sum_{i=1}^{q} p(a_i)\,p(b_j/a_i)}$$

$\left(p(a_i)\,p(b_j/a_i) = p(b_j)\,p(a_i/b_j)\right)$

$p(b_j/a_i)$ = probabilidades directas (hacia adelante)

$p(a_i/b_j)$ = probabilidades inversas (hacia atrás)

5.2.- ENTROPÍA CONDICIONAL

Dado un alfabeto $\{a_i\}$ con probabilidades $\{p(a_i)\}$, podemos calcular la entropía "a priori" como $H(A)$:

$$H(A) = \sum_{i=1}^{q} p(a_i) \log \frac{1}{p(a_i)}$$

Si tenemos los símbolos de salida y sus probabilidades $\{p(a_i/b_j)\}$, se puede calcular la entropía "a posteriori", $H(A/b_j)$:

$$\boxed{H(A/b_j) = \sum_{i=1}^{q} p(a_i/b_j) \log \frac{1}{p(a_i/b_j)}}$$

es la entropía de la v.a. A cuando sé que ha ocurrido b_j

La entropía condicional es la esperanza de $H(A/b_j)$ promediada sobre todos los b_j

$$\boxed{H(A/B) = E\left[H(A/b_j)\right] = \sum_{j=1}^{r} p(b_j)\, H(A/b_j)}$$

mide la incertidumbre de la entrada conociendo la salida, pero sin especificar lo que ha salido del canal

$$H(A/B) = \sum_{j=1}^{r} p(b_j) \sum_{i=1}^{q} p(a_i/b_j) \log \frac{1}{p(a_i/b_j)}$$

$$= \sum_{j=1}^{r} \sum_{i=1}^{q} p(a_i, b_j) \log \frac{1}{p(a_i/b_j)} \quad (*)$$

es una expresión más cerrada

Parece razonable que $H(A) > H(A/B)$, ya que tendrá mayor incertidumbre cuando no se conoce la salida del canal que cuando sí se conoce.

A la entropía condicional también se la llama equivocación del canal.

5.3.- INFORMACIÓN MUTUA

Si $H(A)$ representa la incertidumbre de la entrada y $H(A/B)$ la incertidumbre que queda una vez conocida la salida, la información que se consigue transmitir es:

$$I(A;B) = H(A) - H(A/B) = H(B) - H(B/A) = I(B;A) \quad (*)$$

nos dice lo que disminuye la incertidumbre por conocer la salida

Desarrollando:

$$I(A;B) = \sum_{i=1}^{q} \sum_{j=1}^{r} p(a_i, b_j) \log \frac{p(a_i/b_j)}{p(a_i)}$$

$$= \sum_{j=1}^{r} \sum_{i=1}^{q} p(a_i, b_j) \log \frac{p(a_i, b_j)}{p(a_i)\, p(b_j)}$$

- Propiedades:

a) Conmutativa: $I(A;B) = I(B;A)$

lo que B dice de A es lo mismo que A dice de B

b) Si A y B son variables aleatorias independientes,

$$p(a_i, b_j) = p(a_i)\,p(b_j) \Rightarrow \left\{ \begin{array}{l} H(A/B) = H(A) \\ H(B/A) = H(B) \end{array} \right| \Rightarrow I(A;B) = 0$$

c) $I(A;B) \geq 0 \Rightarrow H(A) \geq H(A/B),\ H(B) \geq H(B/A)$

d) $\min(\log r, \log q) \geq I(A;B)$

$$\left\{ \begin{array}{l} H(B) \leq \log r \\ H(B/A) \geq 0 \end{array} \right. \quad y \quad \left\{ \begin{array}{l} H(A) \leq \log q \\ H(A/B) \geq 0 \end{array} \right.$$

5.4.- ENTROPÍA CONJUNTA

$$S = A \times B = \{ \underbrace{(a_i, b_j)}_{s_k} \}$$

$$p(s_k) = p(a_i, b_j)$$

Suceso simultáneo a_i, b_j ≡ entrada, salida

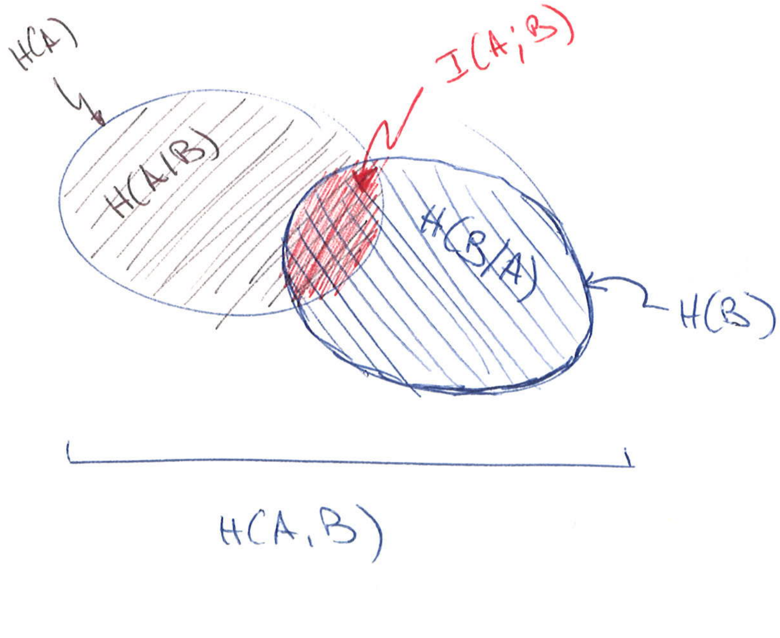
Símbolos s_k a partir de a_i, b_j

$$\boxed{H(S)} = \sum_{k=1}^{q \cdot r} p(s_k) \log \frac{1}{p(s_k)} = \sum_{i=1}^{q} \sum_{j=1}^{r} p(a_i, b_j) \log \frac{1}{p(a_i, b_j)}$$

$$= \boxed{H(A,B)}$$

Con esto, se tiene la última propiedad de la información mutua:

e) $H(A,B) = H(A) + H(B) - I(A;B)$



Ejemplo: Información mutua en un BSC

$$p(b_j/a_i) = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \qquad \begin{matrix} p(a_i=0) = w \\ p(a_i=1) = \bar{w} \end{matrix}$$

$$\begin{cases} p(b_j=0) = \bar{p}\cdot w + p\cdot\bar{w} \\ p(b_j=1) = p w + \bar{p}\,\bar{w} \end{cases} \rightarrow \underline{H(B) = H(wp + \bar{w}\bar{p}) = H(w\bar{p} + \bar{w}p)}$$

$$H(B|A) = \sum_i p(b_j/a_i)\log\frac{1}{p(b_j/a_i)} = \sum_i p(a_i)\underbrace{\sum_j p(b_j/a_i)\log\frac{1}{p(b_j/a_i)}}$$

$H(B/a_i)$, independiente de i, ya que la matriz es simétrica

$$\Rightarrow \sum_i p(a_i) = 1 \Rightarrow \underline{H(B|A) = H(p)}$$

$$\boxed{I(A;B) = H(wp + \bar{w}\bar{p}) - H(p)}$$

Siendo $H(p) = p\log 1/p + (1-p)\log\frac{1}{1-p}$

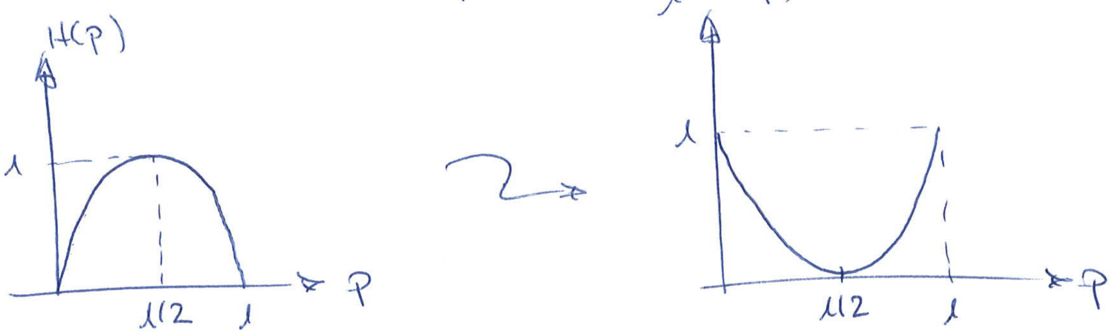
La información entre depende del canal y de las probabilidades de ocurrencia de los símbolos de entrada.

Veamos algunos casos particulares:

a) a_i equiprobables : $w = 1/2$ (caso binario)

$$I(A;B) = H\left(\frac{p+\bar{p}}{2}\right) - H(p) = H(1/2) - H(p)$$

$$= 1 - H(p)$$



- $p = 1/2 \Rightarrow$ máxima incertidumbre ≡ canal pésimo
- $p = 0, 1 \Rightarrow$ el canal no se equivoca nunca o se equivoca siempre, de forma determinista ≡ canal ideal

b) $w = 0$ ó $w = 1$

$$I(A;B) = H(p) - H(p) = 0$$

El conocimiento de la salida no dice nada sobre la entrada, dado que ya se conoce todo sobre la entrada.

c) $p = 1$ ó $p = 0$

$$I(A;B) = H(w) - 0 = H(w)$$

La información mutua es igual a la entropía de la fuente.

Es el canal determinista, en el que el conocimiento de la salida nos dice todo de la entrada

d) $p = 1/2$

$$I(A;B) = H(wp + \bar{w}\bar{p}) - H(p) = H\left(\frac{w+\bar{w}}{2}\right) - H(1/2) =$$

$$= H(1/2) - H(1/2) = 0$$

Cuando en un canal la probabilidad de tener un 0 es igual que la de obtener un 1, cualquiera que sea la entrada, el conocimiento de la salida no dice nada sobre la entrada.

5.5.- CAPACIDAD DE UN CANAL

Es la máxima información mutua que se obtiene probando con todas las distribuciones de probabilidad de la entrada:

$$\boxed{C = \max_{p(a_i)} I(A;B)}$$

Sólo depende de las características del canal

- Propiedades:

a) $\min(\log r, \log q) \geq C \geq 0$

b) $I(A;B)$, función de $p(a_i)$, es continua y sólo presenta un máximo (no tiene máximos locales). Esto simplifica la obtención de ese máximo, que se hará normalmente por métodos numéricos. Sólo en canales muy sencillos es abordable la resolución analítica.

5.6.- CLASIFICACIÓN DE CANALES

(* 4_6 *)

5.6.1.- CANAL SIN RUIDO

Recibido un b_j, se conoce qué a_i lo generó

$$H(A|B) = 0$$

$$I(A;B) = H(A) - H(A|B) = H(A)$$

$$\boxed{C = \max_{p(a_i)} I(A;B) = \log q}$$

Nota: $A = \{a_1, a_2, \dots, a_q\}$

$B = \{b_1, b_2, \dots, b_r\}$

5.6.2.- Canal determinante

Sólo un elemento distinto de cero por cada fila

⇒ dado un a_i, se conoce el b_j de salida

$$\boxed{C = \log r}$$

5.6.3.- Canal uniforme

Todas las filas son permutaciones de la primera.

$$C = \max_{p(a_i)} \{H(B) - H(B/A)\}$$

más fácil calcularlo así que como $H(A) - H(A/B)$

$$H(B/A) = \underbrace{\sum_{i=1}^{q} p(a_i) \sum_j p(b_j/a_i) \log \frac{1}{p(b_j/a_i)}}$$

no depende de i, ya que vale lo mismo para todas las filas

$$\Rightarrow H(B/A) = \sum_j p(b_j/a_i) \log \frac{1}{p(b_j/a_i)}$$

$$\Rightarrow \boxed{C = \max_{p(a_i)} \{H(B)\} - \sum_{j=1}^{r} p(b_j/a_i) \log \frac{1}{p(b_j/a_i)}} \quad (*)$$

Sólo hay que maximizar $H(B)$

$$\max \{H(B)\} \leq \log r$$

La igualdad sólo ocurre cuando hay equiprobabilidad, lo que no puede darse en un canal uniforme (en general)

5.6.4- CANAL DEBILMENTE SIMÉTRICO

Es un canal uniforme en el que se verifica que las sumas de los elementos de cada columna son las mismas.

De esta manera, se garantiza que existe alguna combinación de los símbolos de entrada que hace que los símbolos de salida sean equiprobables.

$$p(b_j) = \sum_{i=1}^{q} p(a_i, b_j) = \sum_{i=1}^{q} p(b_j/a_i)\, p(a_i)$$

Si $p(a_i) = 1/q \quad \forall i$, entonces

$$p(b_j) = \frac{1}{q} \underbrace{\sum_{i=1}^{q} p(b_j/a_i)}_{\text{suma de los elementos de la columna } j}$$

$$\Rightarrow p(b_j) = 1/r = cte$$

Por tanto, su capacidad es

(*)

$$\boxed{C = \log r - \sum_{j=1}^{r} p(b_j/a_i) \log \frac{1}{p(b_j/a_i)}}$$

5.6.5.- CANAL SIMÉTRICO

Es un canal uniforme en el que las columnas son permutaciones arbitrarias de la primera.

¡Ojo! $\neq$ matriz simétrica

Un canal simétrico es también débilmente simétrico, por lo que su capacidad es la misma:

$$\boxed{C = \log r - \sum_{j=1}^{r} p(b_j/a_i) \log \frac{1}{p(b_j/a_i)}}$$

Ejemplo: BSC $2 \Rightarrow P = \begin{bmatrix} p & \bar{p} \\ \bar{p} & p \end{bmatrix}$

$\Rightarrow C = 1 - H(p)$ bits

NOTA: en examen se preguntan capacidades de canales uniformes

Ejemplo: canal uniforme

$$\underline{\underline{P}} = \begin{bmatrix} 1/3 & 1/6 & 1/2 \\ 1/2 & 1/3 & 1/6 \end{bmatrix}$$

$\max\limits_{p(a_i)} \{H(B)\}$?

$A = \{ \underset{p}{a_1}, \underset{\bar{p}}{a_2} \}$ $\qquad B = \{ b_1, b_2, b_3 \}$

$$\begin{cases} p(b_1) = \sum_{i=1}^{2} p(a_i)\, p(b_1/a_i) = p \cdot 1/3 + \bar{p}\, 1/2 \\ p(b_2) = p\, 1/6 + \bar{p}\, 1/3 \\ p(b_3) = p\, 1/2 + \bar{p}\, 1/6 \end{cases}$$

$p(b_1) = p(b_2) = p(b_3)$?

$$p(b_1) = p(b_2) \Rightarrow p(1/3 - 1/6) + \bar{p}\,(1/2 - 1/3) = 0$$

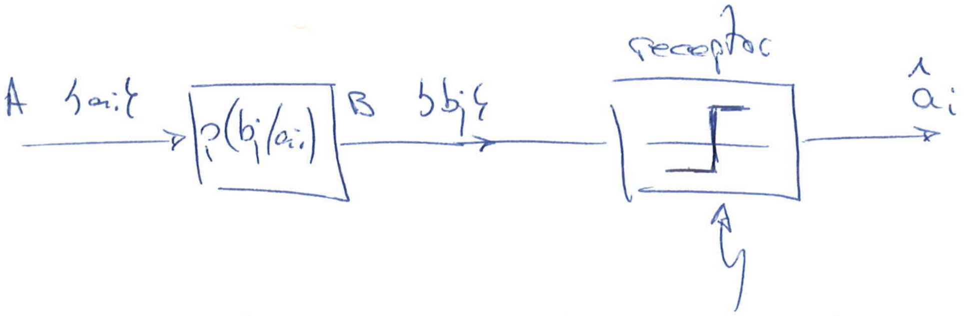
$$p\, 1/3 + (1-p)\, 1/6 = 0 \Rightarrow \underline{p\, 1/6 = -1/6} \;\; !!!$$

No es posible la equiprobabilidad.

$\max \{H(B)\} = 1{,}55$

$\log r = \log 3 = 1{,}59$

5.7.- RECEPTORES: DECISIÓN ÓPTIMA Y ML



<u>decisor</u>: intenta "adivinar" los símbolos transmitidos $a(n)$ a partir de los recibidos $b(n)$

- <u>reglas de decisión</u>:

  - asignar a cada b_j recibido un valor $\hat{a}_i \in A$

    Este $\hat{a}_i$ se denota $d(b_j)$ = "decisión al recibir b_j"

  - Con r símbolos de salida y q símbolos de entrada, se puede hacer $\underline{q^r}$ reglas de decisión posibles

    La mayoría serán absurdas.

- parámetro de calidad del decisor: <u>probabilidad de error</u>

$$P_E = \text{Prob}\{d(b(n)) \neq a(n)\}$$

= probabilidad de que el símbolo recibido sea diferente que el transmitido

· probabilidad de error al recibir b_j:

$$p(E/b_j) = 1 - \underbrace{p(d(b_j)/b_j)}_{\text{probabilidad de acierto}}$$

$$p_E = \sum_j p(E/b_j)\, p(b_j) = \sum_j p(b_j) - \sum_j p(d(b_j)/b_j) \cdot p(b_j)$$

$$\boxed{p_E = 1 - \sum_{j=1}^{s} p(d(b_j), b_j)}$$

La probabilidad de acierto es la suma de las probabilidades de los casos en los que acierto.

5.7.1- REGLA DE DECISIÓN ÓPTIMA

Es aquella que hace mínima la probabilidad de error.

[Nota: nunca decir "más óptima", ya que "óptimo" es un término absoluto]

$$d(b_j) = a^*$$

debe cumplir: $\boxed{p(a^*/b_j) \geq p(a_i/b_j) \ \forall i}$

Se trata de un criterio MAP (Máximo a priori)

El problema de la regla de decisión óptima es que depende de las probabilidades condicionales hacia atrás, $p(a_i/b_j)$, es decir, de la distribución de probabilidades de los símbolos de la entrada.

Se puede expresar con las probabilidades condicionales hacia adelante:

$$\frac{p(b_j/a^*)\cdot p(a^*)}{p(b_j)} \geq \frac{p(b_j/a_i)\cdot p(a_i)}{p(b_j)}$$

$$\boxed{p(b_j/a^*)\, p(a^*) \geq p(b_j/a_i)\, p(a_i) \quad \forall i}$$

5.7.2.- <u>REGLA DE DECISIÓN ML</u> (Maximum Likelihood) (o máxima paridad)

$$\boxed{d(b_j) = a^* \;/\; p(b_j/a^*) \geq p(b_j/a_i) \quad \forall i}$$

Se trata de una particularización de la regla óptima para <u>a_i equiprobables</u>.

En general, el decisor óptimo es mejor que el ML

Ventaja: sólo depende del canal, por ello se usa más que la óptima

Ejemplo:

$$\text{transmisión} \begin{cases} A \longrightarrow A \text{ ó } B & \{ p(A/A) > p(B/A) \\ B \longrightarrow B \text{ ó } A & \{ p(B/B) > p(A/B) \end{cases}$$

(recepción)

Decisor ML:

- si llega A, $\hat{a}_i = A$
- si llega B, $\hat{a}_i = B$

Independientemente de $p(A)$ y $p(B)$ en el transmisor

Si, por ejemplo, es más probable transmitir A que B, puede ser mejor considerar la recepción de B como la transmisión de una A errónea. Esto lo hace el decisor óptimo.

——— o ———

Si los símbolos son equiprobables:

$$P_E = 1 - \sum_i p(d(b_i), b_i) = 1 - \sum_i p(b_i/d(b_i))\underbrace{p(d(b_i))}_{1/q}$$

$$= 1 - \frac{1}{q}\sum_i p(b_i/d(b_i))$$

Ejemplo:

$$\begin{array}{cc} & \longleftarrow b_j \longrightarrow \\ \underline{\underline{P}} = & \begin{bmatrix} 0,5 & 0,3 & 0,2 \\ 0,2 & 0,3 & 0,5 \\ 0,3 & 0,3 & 0,4 \end{bmatrix} \updownarrow a_i \end{array}$$

con a_i equiprobables $\Rightarrow$ óptimo $\equiv$ ML

buscamos $p(b_j/a^*)$ máxima $\Rightarrow$ $\begin{cases} p(b_1/a^*) = 0,5 \\ p(b_2/a^*) = 0,3 \\ p(b_3/a^*) = 0,5 \end{cases}$

$$\begin{cases} d(b_1) = a_1 \\ d(b_2) = a_3 \\ d(b_3) = a_2 \end{cases}$$

Probabilidad de error:

$$\bar{P}_E = \frac{0,5+0,3+0,5}{3} = \frac{1,3}{3} = 0,43\hat{3}$$

$$P_E = 1 - \bar{P}_E = 1 - 0,43\hat{3} = 0,5\hat{6}$$

——— o ———

Ejemplo: BSC

$$\begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \qquad p = 0{,}01$$

$$\bar{p} >> p$$

Si a_i equiprobables, $P_E = 1 - \frac{1}{2}(0{,}99 + 0{,}99) = 0{,}01$

→ de cada 100 bits, uno erróneo

$C = 1 - H(p) = 0{,}92$ → de cada 100 bits, 92 son de info (con o sin error)

Código de repetición: de 3 en 3

$m = 0 \rightarrow c = 000$

$m = 1 \rightarrow c = 111$

$$\Pi \rightarrow \boxed{BSC} \rightarrow \Pi \quad \Rightarrow \quad C \subset \Pi^3 \rightarrow \boxed{BSC} \rightarrow \Pi^3$$

· distancia de Hamming: $v_j \in \Pi^n$

$d_H(v_i, v_j)$ = nº de bits diferentes entre v_i y v_j

$d_H(v_i, v_j)$ = nº de unos de $v_i \oplus v_j$

$0 \to 000 \begin{cases} 1 \text{ error: } 010 \\ 0 \text{ errores: } 000 \end{cases}$

decodificación:

000 → 0		111 → 1	
010 → 0		110 → 1	
100 → 0		101 → 1	
001 → 0		011 → 1	

$$\begin{cases} p(0 \text{ errores}) = \bar{p}^3 & p(2 \text{ errores}) = 3p^2\bar{p} \\ p(1 \text{ error}) = 3p\bar{p}^2 & p(3 \text{ errores}) = p^3 \end{cases}$$

Sólo habrá error en la decodificación si se producen 2 o 3 errores:

$$\boxed{P_e = 3p^2\bar{p} + p^3 \simeq 3 \cdot 10^{-4}}$$

$\underline{R = 1/3}$

Si se usan más bits:

$5 \text{ bits} \to p_e = 10^{-5} \to R = 1/5$

$7 \text{ bits} \to p_e = 4 \cdot 10^{-7} \to R = 1/7$

$9 \text{ bits} \to p_e = 10^{-8} \to R = 1/9$

$\vdots$

$$\boxed{R = \frac{\text{cantidad de información transmitida}}{\text{nº mensajes binarios utilizados}} = \frac{\log M}{n}}$$

Al aumentar el nº de bits, $p_e \downarrow$ y $R \downarrow$

5.8.- Segundo Teorema de Shannon

La probabilidad de error se puede hacer tan pequeña que se desee, manteniendo el factor R tan próximo posible a la capacidad del canal.

$$p_E < \delta \quad y \quad R = C - \varepsilon \quad \text{para} \quad \delta, \varepsilon > 0$$

- <u>Enunciado para canal BSC</u>: sean $\varepsilon, \delta > 0$, valores pequeños, y M el número de símbolos o palabras código para la transmisión $\rightarrow$ $M = 2^{n(C-\varepsilon)}$

Para n suficientemente grande, podemos encontrar un código con M palabras código (alfabeto $A \in \Pi^n$, con $\Pi = \{0, 1\}$), tal que la probabilidad de error sea

$$p_E < \delta$$

y que la información cumple

$$R = \frac{\log M}{n} = C - \varepsilon$$

Es decir, aumentando n se puede conseguir $\begin{cases} P_E \to 0 \\ R \to C \end{cases}$

y para ello se puede obtener una codificación que lo permita.

Codificación de canal en general:

$$\Pi \longrightarrow \boxed{BSC} \longrightarrow \Pi \qquad\qquad \underline{P}$$

$$\Pi^3 \longrightarrow \boxed{BSC} \longrightarrow \Pi^3 \qquad\qquad 2^n \updownarrow \underbrace{\left[\quad \underline{P}_n \quad\right]}_{2^n}$$

$$C' \subset \Pi^n \longrightarrow \boxed{BSC} \longrightarrow \Pi^n \qquad\qquad M \updownarrow \underbrace{\left[\qquad\quad\right]}_{2^n}$$

(M = nº palabras código para la transmisión)

Dados M y n, hay que definir

C' tal que P_e sea mínima.

Se definirá basándonos en la distancia de Hamming

Shannon demostró cómo de bien se puede codificar el canal:

a) para $\Pi \Rightarrow R = \dfrac{\log_2(2^n)}{n} = 1$

b) para M $\Rightarrow$ $R = \frac{\log M}{n} \leq 1$

$$\boxed{C = 1 - H(p)}$$ p: probabilidad de error del canal (BSC)

Enunciado del <u>segundo teorema de Shannon:</u> (canal BSC)

$\forall\, \epsilon, \delta > 0$, $\exists\, n$ suficientemente grande /

$\exists\, C, M = 2^{n(C-\epsilon)}$ / $p_e < \delta$

y tal que la información cumple:

$$R = \frac{n(C-\epsilon)}{n} = C - \epsilon$$

- asintóticamente, la probabilidad de error es muy baja con el rendimiento que se quiera

$H(M) = \log_2(M)$ $\Rightarrow$ cantidad de información por símbolo del canal

La <u>máxima cantidad de información</u> por símbolo es igual a la capacidad.

Ejemplo: BSC $\rightarrow C = 1 - H(p)$

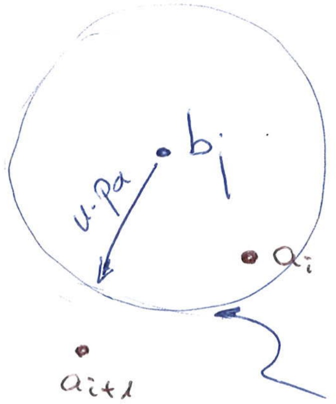
- número medio de bits erróneos en una secuencia de n bits: $n \cdot p$

- regla de decisión en recepción: aquella palabra cuya distancia d_H sea menor que $n \cdot p$

$\rightarrow$ más exacto: $d_H < n(p+\alpha)$

con $\alpha > 0$ arbitrariamente pequeño

$d_H < n \cdot p\alpha$



$b_j \rightarrow$ recibido

$a_i \rightarrow$ se selecciona el a_i cuya distancia con b_j sea menor a $n \cdot p\alpha$

esfera $\equiv S(n \cdot p\alpha)$

La probabilidad de error será:

$P_E \equiv$ probabilidad de que el a_0 original que provocó b_j no esté dentro de la esfera

$+$

probabilidad de que el a_0 esté dentro pero haya otro a_K dentro con él

$$p_E = p(a_0 \notin S(np\alpha)) + p(a_0 \in S(np\alpha))\, p(a_{i=0} \in S(np\alpha))$$

$$p_E < p(a_0 \notin S(np\alpha)) + \underbrace{p(a_{i=0} \in S(np\alpha))}_{< \sum_{a_i \neq a_0} p(a_i \in S(np\alpha))}$$

$$p_E < p(a_0 \notin S(np\alpha)) + \sum_{a_i \neq a_0} p(a_i \in S(np\alpha))$$

<u>Teorema de los grandes números:</u>

al aumentar el número de bits (n), la probabilidad de error se aproximará a la media

$$n\uparrow \Rightarrow p(a_0 \notin S(np\alpha)) < \delta$$

Para el segundo término: $p(a_i \in S(np\alpha)) = \dfrac{N(np\alpha)}{2^n}$

$N(np\alpha)$ = nº de palabras código dentro de la esfera

2^n = número de palabras

$$p_E < \delta + (M-1)\frac{N(np\alpha)}{2^n} < \delta + M\frac{N(np\alpha)}{2^n}$$

Nº de palabras código a una distancia dada k: nº de números posibles de cambiar los bits k veces

$\Rightarrow d_H = k$; nº palabras $= \binom{n}{k}$

Entonces: $N(np\alpha) = \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{np\alpha} = \sum_{k=0}^{np\alpha} \binom{n}{k} \leq 2^{nH(p\alpha)}$

porque sí

$$p_E < \delta + M\frac{2^{nH(p\alpha)}}{2^n} = \delta + M \cdot 2^{-n(1-H(p\alpha))}$$

Si $M = 2^{n(C-\varepsilon)}$

$C = 1 - H(p)$

$$\Rightarrow p_E < \delta + 2^{n(C-\varepsilon)} 2^{-n(1-H(p\alpha))} = \delta + 2^{-n(H(p)+\varepsilon-H(p\alpha))}$$

$\delta \simeq 0$

Si es $H(p)+\varepsilon H(p\alpha) > 0$, entonces $2^{-n\cdot(\dots)} \longrightarrow 0$ con $n\uparrow$

$$H(p) + \varepsilon - H(p\alpha) > 0$$

$$\Downarrow$$

$$\simeq 0 \qquad \simeq 0$$

$$H(p) + \varepsilon - (H(p) + H'(p)\alpha) > 0 \Rightarrow \varepsilon > H'(p)\alpha$$

(fin demostración)

- Interpretación del 2º teorema de Shannon:

$$\underset{H(S)}{\boxed{\text{fuente}}} \xrightarrow{s(u)\in S} \boxed{\text{codificador}} \xrightarrow{b_u \in B} \underset{C}{\boxed{\text{canal}}} \xrightarrow{b'_u \in B} \boxed{\text{decodificador}} \xrightarrow{\hat{s}(u)\in S} \boxed{\text{destino}}$$

velocidad de la fuente: $\boxed{V_F = \dfrac{H(S)}{T_F} \ [bps]}$

T_F = periodo de símbolo

Existen un codificador y decodificador tales que

$p_E < \delta$ y $V_F = \dfrac{H(S)}{T_F} \longrightarrow \dfrac{C}{T_S}$

Seña : $V_F = \dfrac{H(S)}{T_F} = \dfrac{C}{T_C} - \varepsilon$

6.- Teoría de la Información para Variables Aleatorias Continuas

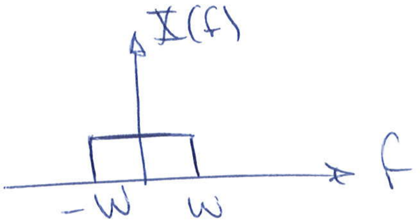
$$s(n) \longrightarrow \boxed{t_x} \xrightarrow{x(t)} \boxed{\text{canal}} \xrightarrow{y(t)} \boxed{r_x} \longrightarrow s'(n)$$

estudiaremos las características del medio analógico

Queremos determinar la capacidad del medio:

$$x(t) \longrightarrow \boxed{\text{medio}} \longrightarrow y(t)$$

$x(t)$ limitada en banda $\Rightarrow$ se puede discretizar en el tiempo

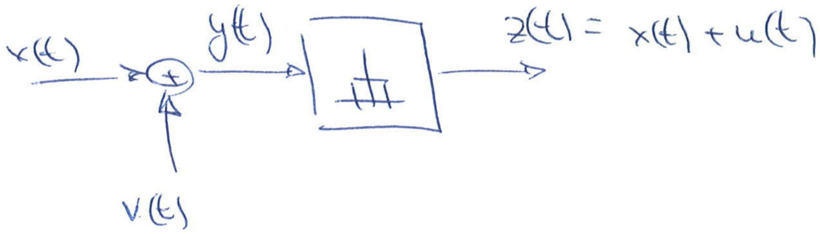


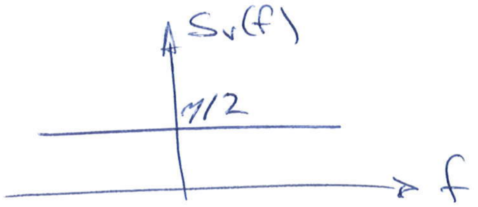
Sistema LTI $\Rightarrow$ $y(t)$ limitada en banda $\Rightarrow$ también se puede discretizar

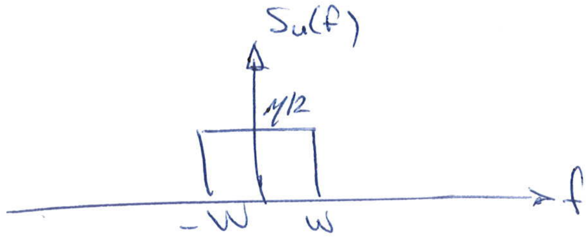
$(f_s \geq 2W)$ $\qquad y(t) \longrightarrow \boxed{} \longrightarrow z(t)$

$$\begin{cases} x(t) \longrightarrow x[n] = x(nT_s) = x\left(\frac{n}{2W}\right) \\ z(t) \longrightarrow z[n] = z(nT_s) = z\left(\frac{n}{2W}\right) \end{cases}$$

Si el medio sólo añade AWGN:

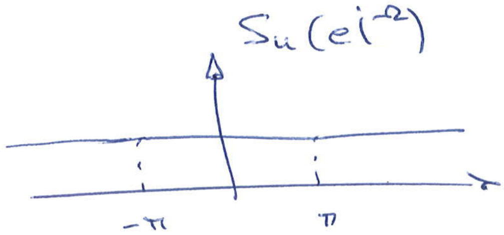






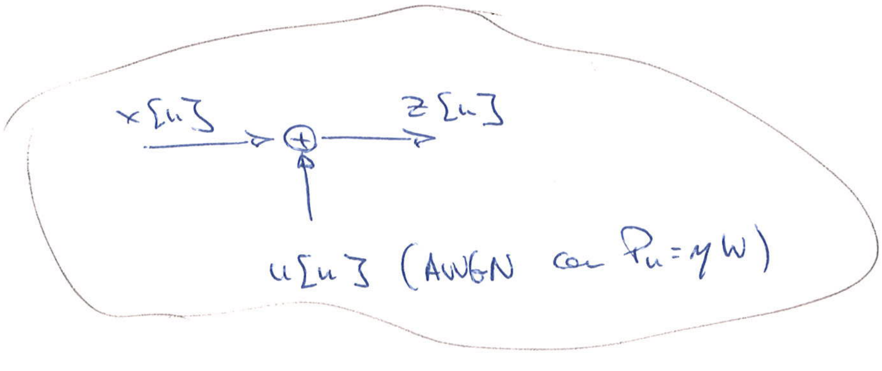
$P_u = \eta W$

Si se muestrea $u(t)$:



ruido blanco y gaussiano

$$R_u[n] = \delta[n] P_u = \eta W \delta[n]$$



6.1.- LEY DE SHANNON-HARTLEY

Establece la capacidad del canal:

(*)

$$\boxed{C = \frac{1}{2}\log\left(1+\frac{P_x}{P_u}\right) = \frac{1}{2}\log\left(1+\frac{S}{N}\right)} \quad [\text{bits/muestra}]$$

$P_x = E[x^2[n]]$ $\qquad T_m = 1/2W$ (intervalo de muestreo)

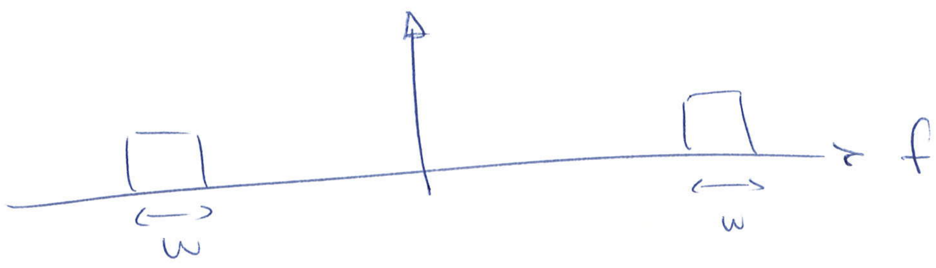
$$\frac{C}{T_m}\,[\text{bps}] = W\cdot\log\left(1+\frac{P_x}{P_u}\right) = W\cdot\log\left(1+\frac{S}{N}\right)\ [\text{bps}]$$

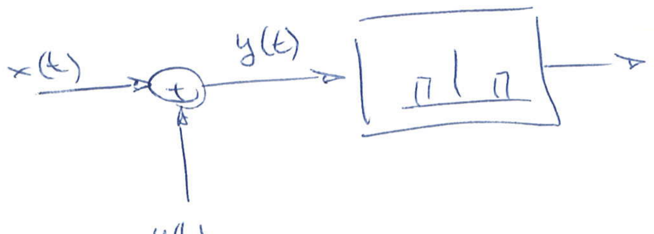
$W \rightarrow [\text{Hz}]$

$$\boxed{C = W\log\left(1+\frac{P_x}{P_u}\right)} \quad [\text{bps}]$$

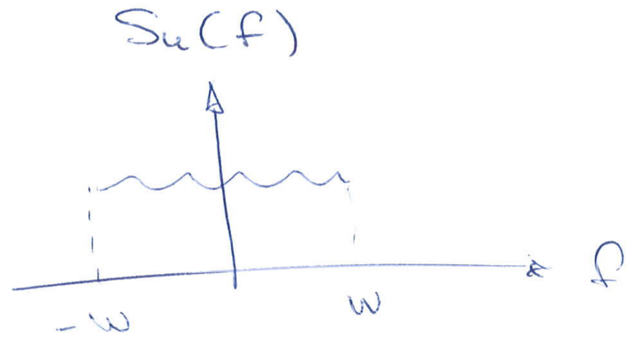
$P_u = \eta W$
(ruido blanco)

Esta capacidad es extrapolable a canales paso banda:





Si suponemos ahora que el ruido es gaussiano pero no blanco (coloreado, AGN):



Canal paso bajo de ancho de banda W

Nos dan la DEP de la señal, $S_x(f)$

Se calcula la capacidad en tramos diferenciales $\Rightarrow$ integrar:

$$dC = df \cdot \log\left(1 + \frac{S_x(f_0)\, 2 \cdot df}{S_u(f_0)\, 2 \cdot df}\right)$$

$$\Rightarrow \boxed{C = \int_0^W \log\left(1 + \frac{S_x(f)}{S_u(f)}\right) df} \quad (*)$$

Ley de Shannon-Hartley para ruido coloreado

- Se fija la potencia de la señal en vez de su DEP

Normalmente: • transmisor con P_x

• canal con $S_v(f)$

$\Rightarrow$ podemos distribuir P_x de muchas formas, verificándose

$$\boxed{P_x = \int_{-W}^{W} S_x(f)\,df}$$

6.2.- LEY DE VERTIDO DE AGUA

(AKA "water-pouring")

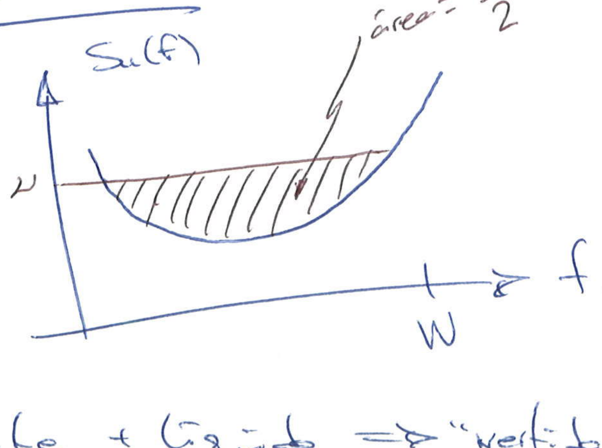
Da la expresión de la DEP óptima de señal que maximiza la capacidad del canal:

$$\boxed{S_x(f) = [\nu - S_n(f)]^+}$$

se tiene como dato la DEP del ruido

siendo $[x]^+ = \begin{cases} x & x \geq 0 \\ 0 & \text{resto} \end{cases}$

Gráficamente:

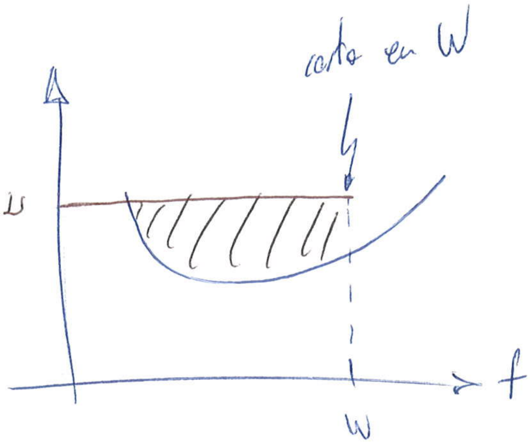


condición para calcular ν

$$\int_0^W S_x(f)\,df = \frac{P_x}{2} \Rightarrow \nu$$

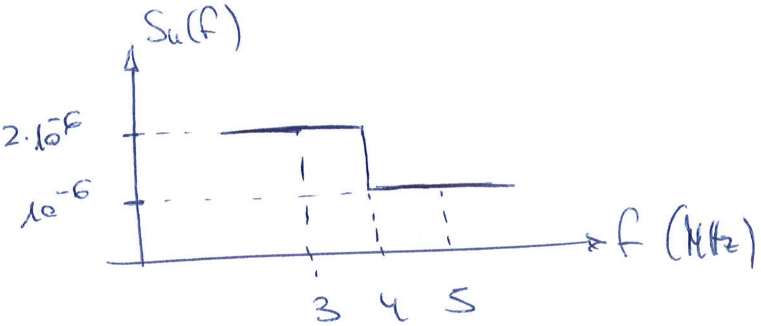
recipiente + líquido $\Rightarrow$ "vertido de agua"

• donde hay poco ruido se pone mucha potencia de señal, y al revés



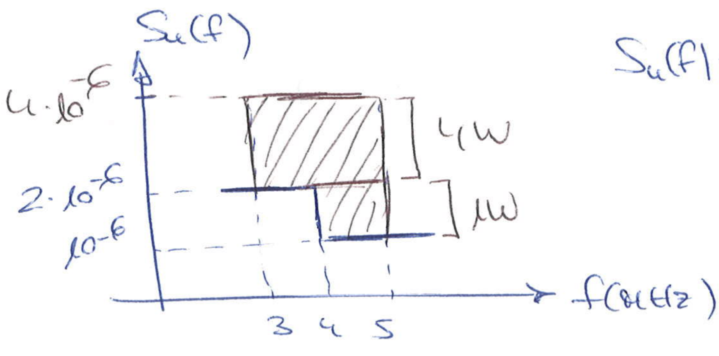
Una vez calculado $S_x(f)$, se aplica la forma integral de la Ley de Shannon-Hartley

Ejemplo:

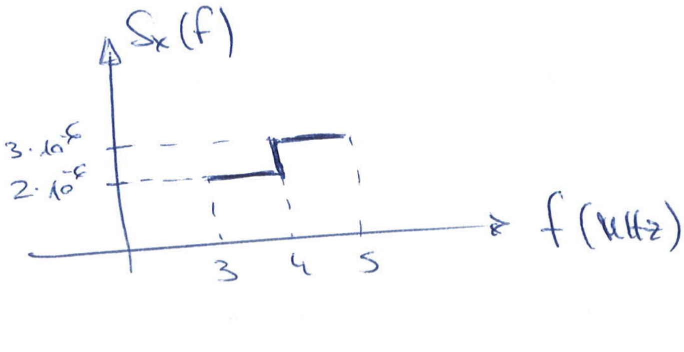


ver la capacidad en la banda entre 3 y 5 MHz

Si $P_x = 10\,W$:



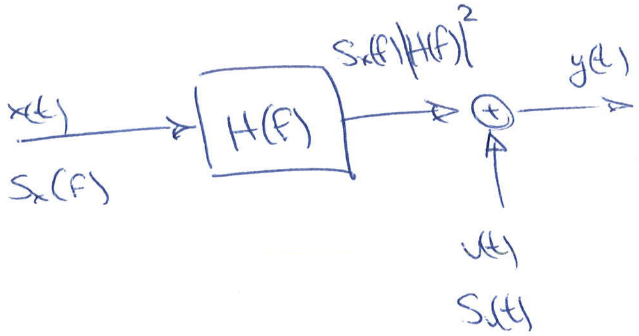
$$S_u(f) = \begin{cases} 2\cdot 10^{-6}\ W/Hz & 3<|f|<4 \\ 3\cdot 10^{-6}\ W/Hz & 4<|f|<5 \end{cases}$$



$$C = 10^6 \left(\log(1+1) + \log(1+3) \right) = 3 \cdot 10^6 \text{ bps}$$

———— o ————

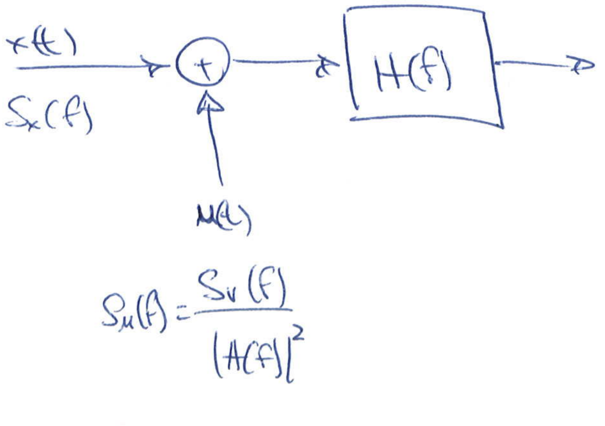
Modelo del canal con distorsión de amplitud y ruido coloreado:



$$\boxed{C = \int_0^W \log\left(1 + \frac{|H(f)|^2 S_x(f)}{S_v(f)}\right) df} \quad [\text{bps}]$$

Si se conocen P_x y $S_v(f)$ se modifica el modelo:

el filtro no afecta a la capacidad del canal



$$S_x(f) = \left[\nu - \frac{S_v(f)}{|H(f)|^2} \right]^+$$

Si se usa esa distribución, la capacidad máxima es:

$$C = \int_0^W \log\left(1 + \frac{\mu - \frac{S_v(f)}{|H(f)|^2}}{\frac{S_v(f)}{|H(f)|^2}}\right) df$$

- El ruido gaussiano es el peor, al ser el más aleatorio
- Si es blanco, es aún peor.

6.3.- ENTROPÍA DIFERENCIAL

Hasta ahora hemos trabajado con señales de tiempo continuo que se pueden discretizar

Las variables aleatorias son continuas, por lo que pueden tener infinitos valores: $N = \infty \Rightarrow \log N = \infty$

$\Rightarrow$ la entropía puede hacerse infinita

Por ello, no tiene sentido aplicar la entropía tal y como la hemos definido a variables aleatorias continuas

Dada una variable aleatoria x con fdp f(x),

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{a}^{b} f(x)\,dx = P[a \le x \le b]$$

Se define su <u>ENTROPÍA DIFERENCIAL</u> como:

$$\boxed{h(X) = -\int_{-\infty}^{\infty} f(x) \log f(x)\,dx}$$

(mismas unidades que en el caso discreto)

- <u>Variable aleatoria uniforme en [0,a]</u>:

$$f(x) = \frac{1}{a} \qquad x \in [0,a]$$

$$\boxed{h(X) = -\int_{0}^{a} \frac{1}{a} \log \frac{1}{a}\,dx = \log a} \quad [\text{bits}]$$

Esto se puede extender para intervalos $[\Delta, \Delta+a]$

Si <u>$a<1$</u>, la entropía diferencial es negativa

- Variable aleatoria gaussiana:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{x^2}{2\sigma^2}} \qquad \bar{x} = 0$$

$$h(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{x^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi}\,\sigma}\, e^{\frac{-x^2}{2\sigma^2}}\right) dx = \{\ldots\} =$$

$$= \frac{1}{2}\log 2\pi\sigma^2 \underbrace{\int_{-\infty}^{\infty} f_x(x)\,dx}_{1} + \int_{-\infty}^{\infty} f_x(x)\,\frac{x^2}{2\sigma^2}\log e\, dx =$$

$$= \frac{1}{2}\log 2\pi\sigma^2 + \frac{\log e}{2\sigma^2}\underbrace{\int_{-\infty}^{\infty} x^2 f_x(x)\,dx}_{\sigma^2}$$

(*)

$$\boxed{h(X) = \frac{1}{2}\log 2\pi e\sigma^2} \text{ [bits]}$$

Este resultado es insensible a los desplazamientos, no depende de la media:

$$h(X) = h(X + c)$$

→ la entropía diferencial es independiente de la media de la variable aleatoria

- ENTROPÍA DIFERENCIAL CONJUNTA

2 variables aleatorias: $X, Y \rightarrow f_x(x), f_y(y), f_{xy}(x,y)$

$$\boxed{h(X,Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,y)\log\left(f_{xy}(x,y)\right)dx\,dy}$$

- ENTROPÍA DIFERENCIAL CONDICIONAL

$$\boxed{h(X/Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,y)\log\left(f_{xy}(x/y)\right)dx\,dy}$$

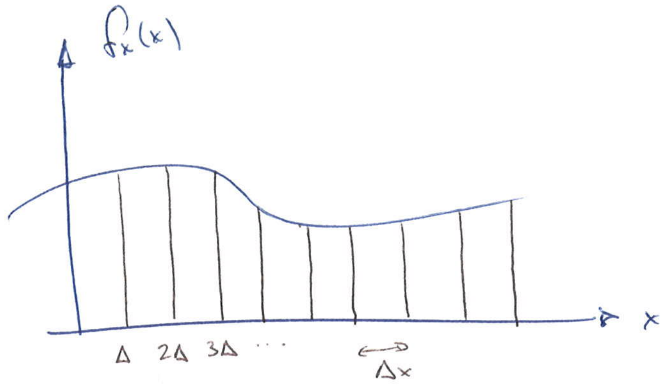
- INFORMACIÓN MUTUA DIFERENCIAL

$$\boxed{I(X;Y) = h(X) - h(X/Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,y)\log\frac{f_{xy}(x,y)}{f_x(x)\,f_y(y)}\,dx\,dy}$$

$$I(X;Y) = I(Y;X)$$

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

No hay relación entre las definiciones de entropía y entropía diferencial, pero sí entre las informaciones mutuas de variables aleatorias discretas y continuas.



Variables aleatorias:

- continua: X, x
- discreta: X^Δ, x_i

Intentamos calcular la entropía "convencional" construyendo la variable aleatoria discreta X^Δ, segmentando X con intervalos de longitud Δ:

$$X_i \in X^\Delta \longrightarrow i\Delta < X_i \leq (i+1)\Delta$$

$$P_i = \int_{i\Delta}^{(i+1)\Delta} f_X(x)\,dx = f_X(x_i)\cdot\Delta$$ (teorema del valor medio)

$x_i \equiv$ punto del valor medio

Calculando su entropía:

$$H(X^\Delta) = -\sum_i P_i \log P_i = -\sum_i f_X(x_i)\,\Delta\left(\log f_X(x_i) + \log \Delta\right) =$$

$$= -\sum_i f_X(x_i)\,\Delta \log f_X(x_i) - \log \Delta$$

Calculando el límite en $\Delta \to 0$:

$$\lim_{\Delta \to 0} H(X^\Delta) = h(X) - \log \Delta$$

$$h(X) = \lim_{\Delta \to 0} H(X^\Delta) + \lim_{\Delta \to 0} \log \Delta = \infty - \infty$$

Por tanto, no hay relación.

Sin embargo, sí que la hay para la información mutua:

$$I(X^\Delta ; Y^\Delta) = H(X^\Delta) - H(X^\Delta / Y^\Delta) \overset{\Delta \to 0}{\simeq} \quad \text{(igualdad en el límite)}$$

$$\simeq h(X) - \cancel{\log \Delta} - h(X/Y) + \cancel{\log \Delta}$$

$$= h(X) - h(X/Y)$$

$$\boxed{I(X ; Y) = \lim_{\Delta \to 0} I(X^\Delta ; Y^\Delta)}$$

- Propiedades:

1.- $I(X ; Y) \geq 0$ es nula cuando X e Y son independientes

2.- $h(X/Y) \leq h(X)$

el conocimiento de la variable Y sólo puede incrementar el conocimiento de X

3.- $h(X + c) = h(X)$ con c constante

4.- $h(a \cdot X) = h(X) + \log |a|$

6.4.- DISTRIBUCIONES DE MÁXIMA ENTROPÍA DIFERENCIAL

a) Valores en $(-\infty, \infty)$, para potencia restringida: $\overline{x^2} = \sigma^2$

$$E[x^2] = \sigma^2 \text{ constante} \quad , \quad E[x] = 0$$

no es necesario indicarlo, ya que la media no influye en la entropía

En general: $f(x)$, $E[(x-m_x)^2] = \sigma^2$

Distribución gaussiana: es la que maximiza la entropía

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

b) Valores en $[0, \infty)$, para una media restringida:

$$E[x] = m$$

Distribución exponencial: $f_x(x) = \frac{1}{m} e^{-mx} \quad x \geq 0$

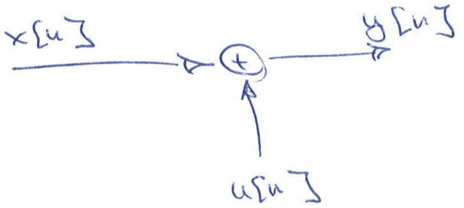
c) Valores en $[0,A]$ (o en $[a,b]$)

Distribución uniforme: $f(x) = \frac{1}{A} \quad x \in [0,A]$

En condiciones en que sólo fijamos la potencia, la distribución que más entropía (incertidumbre) tiene, es la gaussiana, por lo que es más perniciosa en cuanto al ruido.

6.5.- CAPACIDAD DE UN CANAL GAUSSIANO

Volvemos al canal para justificar la ley de Shannon-Hartley



queremos calcular la capacidad $\Rightarrow$ información mutua

$u[n]$ v.c. gaussiana, potencia σ_u^2

$$I(X;Y) = h(Y) - h(Y/X) = h(Y) - h\left(X + U/X\right)$$

si presuponemos que conocemos X, la única variable que tiene aleatoriedad es U

$$h(X+U/U) = h(X/X) + h(U/X) \; ; \; h(X/X)=0$$

$$\Longrightarrow h(Y) - h(U/X) = h(Y) - h(U)$$

$$h(U) = \frac{1}{2}\log\left(2\pi e \sigma_u^2\right)$$

$$C = \max_{f_x(x)} I(X;Y)$$

si no se ponen más restricciones, se puede hacer infinita

- restringir la potencia de X: $E[x^2] \leq P_x$

$$\Longrightarrow E[y^2] = P_x + \sigma_u^2$$

Para que y sea gaussiana, x tiene que ser gaussiana:

$$\max h(Y) = \frac{1}{2}\log\left(2\pi e \left(P_x + \sigma_u^2\right)\right)$$

$$C = \max_{\substack{f(x) \\ E[x] \leq P_x}} I(X;Y) = \frac{1}{2}\log\left(2\pi e\left(P_x + \sigma_u^2\right)\right) - \frac{1}{2}\log\left(2\pi e \sigma_u^2\right)$$

$$\Longrightarrow \boxed{C = \frac{1}{2}\log\left(1 + \frac{P_x}{\sigma_u^2}\right)}$$

Canal gaussiano:

"Elements of Information Theory", p 240 →

http://cnx.org/content/m0073/latest

BSC ⟹ $\boxed{C = 1 + p_e \log_2 p_e + (1-p_e)\log_2(1-p_e)}$

Canal gaussiano a esquema de decodificação

$+\sqrt{P}, -\sqrt{P}$ $\qquad y \geq 0 \Rightarrow +\sqrt{P}$

$y < 0 \Rightarrow -\sqrt{P}$

$$\Rightarrow p_e = 1 - Q\left(\frac{\sqrt{P}}{N}\right)$$

$$Q(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

Comunicaciones Analógicas

- Tema 1

· entropía: $H(S) = E[I(S_i)] = -\sum_{i=1}^{N} p_i \log p_i \leq \log N$ [bits]

· velocidad de la fuente: $V(S) = \frac{H(S)}{T}$ [bps]

· extensión de la fuente: $H(S^n) = n \cdot H(S)$ $V = \frac{H(S)}{T}$

· inecuación de Kraft-McMillan: $\boxed{\sum_{i=1}^{q} r^{-l_i} \leq 1}$ r = nº símbolos alfabeto código
l_i = longitudes palabras código

· longitud media de un código: $\boxed{L = \sum_{i=1}^{q} p_i l_i}$ $H_r(S) \leq L$

· primer teorema de Shannon: $\boxed{\lim_{n \to \infty} \frac{L_n}{n} = H_r(S)}$

· rendimiento de un código: $\boxed{\eta = \frac{H_r(S)}{L}}$

· entropía condicional: $\boxed{H(A/B) = E[H(A/b_j)] = \sum_{j=1}^{r} p(b_j) H(A/b_j)}$

$$\boxed{H(A/b_j) = \sum_{i=1}^{q} p(a_i/b_j) \log \frac{1}{p(a_i/b_j)}}$$

· información mutua: $I(A;B) = H(A) - H(A/B) = H(B) - H(B/A) = I(B;A)$

· entropía conjunta: $H(A,B) = \sum_{i=1}^{q} \sum_{j=1}^{r} p(a_i, b_j) \log \frac{1}{p(a_i, b_j)}$

· capacidad de un canal: $C = \max_{p(a_i)} \{ I(A;B) \}$

→ canal sin ruido: $C = \log q$

→ canal determinante: $C = \log r$

→ canal uniforme: $C = \max_{p(a_i)} \{H(B)\} - \sum_{j=1}^{r} P(b_j/a_i) \log \frac{1}{p(b_j/a_i)}$

→ canal simétrico/débilmente simétrico:

$$C = \log r - \sum_{j=1}^{r} p(b_j/a_i) \log \frac{1}{p(b_j/a_i)}$$

- decisión: $p_e = 1 - \sum_{j=1} p(d(b_j), b_j)$

→ regla óptima: $p(a^*/b_j) \geq p(a_i/b_j) \quad \forall i$

$$\boxed{p(b_j/a^*)\, p(a^*) \geq p(b_j/a_i)\, p(a_i) \quad \forall i}$$

→ regla ML: $\boxed{d(b_j) = a^* \;/\; p(b_j/a^*) \geq p(b_j/a_i) \quad \forall i}$

- segundo teorema de Shannon:

$\forall \varepsilon, \delta > 0, \ \exists n$ (grande) $/ \ \exists \mathcal{C}, \ M = 2^{n(C-\varepsilon)} \ / \ p_e < \delta$

y la información cumple:

$$R = \frac{n(C-\varepsilon)}{n} = C - \varepsilon$$

$\mathcal{C}$: código
C: capacidad
$R = \frac{\log M}{n}$

- ley de Shannon-Hartley: $C = \frac{1}{2}\log\left(1+\frac{S}{N}\right)$ [bits/muestra]

$$= W \log\left(1+\frac{S}{N}\right) \quad [\text{bps}]$$

→ ruido coloreado: $\boxed{C = \int_0^W \log\left(1+\frac{S_x(f)}{S_u(f)}\right) df}$

- ley de vertido de agua: $S_x(f) = [\nu - S_u(f)]^+$

$$\int_0^W S_x(f)\,df = \frac{P_x}{2} \Rightarrow \nu$$

- ley de Shannon-Hartley con distorsión de amplitud:

$$\boxed{C = \int_0^W \log\left(1+\frac{|H(f)|^2 S_x(f)}{S_v(f)}\right) df}$$

$$S_v(f) = |H(f)|^2 S_u(f) \qquad S_x(f) = \left[\nu - \frac{S_v(f)}{|H(f)|^2}\right]^+$$

$$C = \int_0^W \log\left(1+\frac{\nu - \frac{S_v(f)}{|H(f)|^2}}{\frac{S_v(f)}{|H(f)|^2}}\right) df$$

- entropía diferencial: $\boxed{h(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)\,dx}$

- V.A. uniforme en $[0,a]$: $h(X) = \log a$
- V.A. gaussiana: $h(X) = \frac{1}{2}\log 2\pi e\sigma^2$

- entropía diferencial conjunta:

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

- entropía diferencial condicional:

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

- información mutua diferencial:

$$I(X;Y) = h(X) - h(X/Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log \frac{f(x,y)}{f(x)\,f(y)}\, dx\, dy$$

- capacidad de un canal gaussiano:

$$C = \frac{1}{2} \log\left(1 + \frac{P_x}{\sigma_n^2}\right)$$