Introducción a la Teoría de la Información

Prueba final

9 de julio de 2018

Problema 1 (10 puntos)

El *tiro desde los once pasos* es la pena máxima en el fútbol; se desea analizar con herramientas de teoría de la información las opciones del guardameta. El arco de fútbol se divide en cinco zonas $Z = \{z_1, z_2, z_3, z_4, z_5\}$ hacia donde el golero puede arrojarse. Según el perfil del ejecutante (zurdo o diestro), el arquero tiene una distribución de probabilidad $p_I(Z)$ o $p_D(Z)$ de las zonas como se muestra en la siguiente figura.

1. Hallar la información media que tiene el cuidavallas si sabe el perfil del ejecutante, tanto para el caso que sea zurdo o diestro.
2. Hallar un código de largo medio mínimo para codificar las zonas del arco en ambos casos.
3. Verificar que los código hallados son óptimos para las distribuciones respectivas.

Considerar que la probabilidad de que el ejecutante sea zurdo es $1/10$. Para codificar la zona hacia donde se arrojó el portero se propone utilizar un bit para codificar el perfil del ejecutante y luego concatenar con el código de la región correspondiente según dicho perfil.

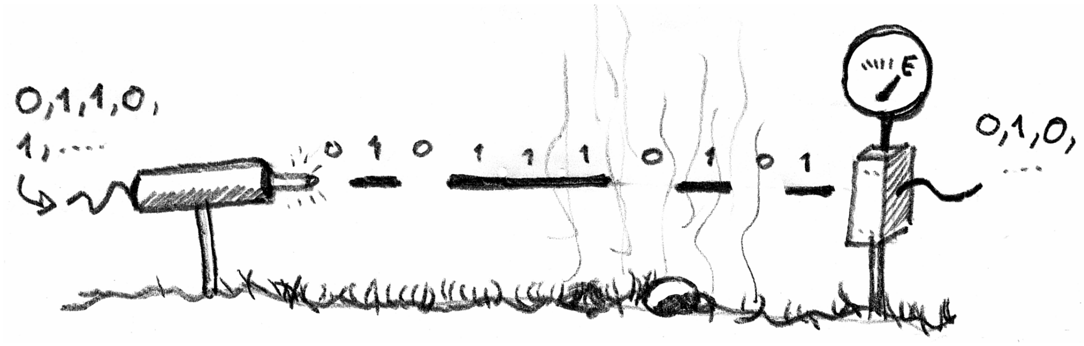
4. Hallar el largo medio de esta codificación así generada.
5. Existe una forma más eficiente, ¿cuál es?

Problema 2 (15 puntos)

Un canal transmite una variable discreta binaria, $X = V$ o $X = -V$ con probabilidades $1-p$ y p respectivamente. El canal tiene ruido Z de distribución uniforme en $(-a, a)$. Se recibe $Y = X + Z$.

1. Expresar o dibujar la densidad de probabilidad de Y distinguiendo los casos: (1) $a < V$ y (2) $V < a < 2V$.
2. Hallar la entropía discreta o continua, según corresponda, de X, Z e Y, en los dos casos citados. Hallar la entropía condicional $H(Y|X)$. Comentar el resultado $H(Y)$ en el caso 1.
3. Hallar la capacidad del canal en los dos casos. Comentar cómo varía con a.
4. Se decide si se recibió V o $-V$ comparando Y con un umbral U. ¿Cuál es el mejor valor de U para minimizar la probabilidad de error en cada uno de los casos 1 y 2?

Problema 3 (15+2 puntos)



Considere un sistema de transmisión por láser como el de la figura. El transmisor envía la señal como pulsos de luz a intervalos regulares de 1 ns (nanosegundo = 10^{-9} segundos). En este canal binario, los 1 son enviados manteniendo el pulso encendido durante todo el intervalo. Esto equivale a una cierta energía E J (Joules).
En cada intervalo, el receptor decide haber recibido un 1 si durante ese tiempo la energía recibida es al menos $E/2$ J; por debajo de ese umbral, se declara haber recibido un 0.
Normalmente, el pulso de energía asociado a un 1 es recibido sin pérdida alguna, es decir, el receptor recibe exactamente E J por cada 1, y 0 J por cada cero enviado.
Sin embargo, debido a la turbulencia atmosférica, se dan intervalos de tiempo durante los cuales el láser es desviado y no arriba al detector. Estos intervalos ocurren en promedio $\tau = 0{,}001$ vez por ns (es decir, uno de cada 1000 muestras), y tienen una duración con distribución uniforme entre 0 y β ns.
Si se produce un desvío durante la transmisión de un 1, se perderá una cantidad de energía proporcional a cuánto de ese desvío cae dentro del tiempo de transmisión del bit (1 ns). Para que un bit 1 sea recibido como 0, el desvío debe afectar más del 50 % del tiempo de bit.
Supongamos de aquí en adelante que la probabilidad de que dos desvíos consecutivos ocurran en menos de 1 ns es despreciable. Tenemos tres casos:

a) $\beta < 0{,}5$ ns

b) $0{,}5 \leq \beta < 1$ ns

c) $1 < \beta < 2$ ns

Para cada caso se pide:

1. Dibujar la representación gráfica del canal (grafo bipartito).

2. Indicar si el canal descrito se parece o es idéntico a algún canal visto en el curso.

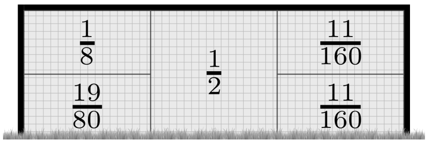
3. Suponiendo que el canal del caso b) no tiene memoria y su probabilidad de error de bit es $\alpha = 1/2$ (nada de esto es necesariamente cierto), calcule la capacidad del canal correspondiente y la distribución de la entrada que la alcanza.

4. **(2 Puntos extra)** ¿En caso de que uno (o más) de los canales tengan memoria, se le ocurre alguna manera de aprovechar esa memoria para mejorar la transmisión? Describa su estrategia en pocas en palabras.

Solución problema 1

1. $H_D(Z) = 15/8; H_I(Z) = 2$
2. El código de largo medio mínimo se puede obtener con Huffman. Los códigos de Huffman para ambos casos quedan: $C_D = 1, 10, 110, 1110, 1111$ $C_I = 1, 000, 001, 010, 011$
3. $L_I = H_I(Z); L_D = H_D(Z)$
4. $\tilde{L} = 1 + p_I H_I + p_D H_D = 1 + \frac{1}{10} \times 2 + \frac{9}{10} \times \frac{15}{8} = \frac{231}{80} = 2{,}8875$
5. Combinando las probabilidades de ambos perfiles, tenemos una nueva distribución

$$p_Z = \{1/2, 1/8, 19/80, 11/160, 11/160\}.$$



El código de Huffman para esta distribución queda

$$C_Z = \{0, 100, 11, 1010, 1011\}$$

y el largo medio 1.9, siendo más eficiente que la codificación anterior.

Solución problema 2

Z es uniforme en el intervalo $(-a, a)$ con densidad constante $\frac{1}{2a}$.

1. En el caso 1 la densidad de probabilidad de Y es $(1-p)\frac{1}{2a}$ en el intervalo $(-V-a, -V+a)$ y $p\frac{1}{2a}$ en $(V-a, V+a)$. Los intervalos son disjuntos.

   En el caso 2 hay superposición. La densidad de Y es $(1-p)\frac{1}{2a}$ en el intervalo $(-V-a, V-a)$, $\frac{1}{2a}$ en $(V-a, -V+a)$ y $p\frac{1}{2a}$ en $(-V+a, V+a)$.
2. $H(X) = h(p)$, la función entropía.

   $H(Z) = \log 2a$, entropía continua.

   $H(Y|X) = H(Z) = \log 2a$

   Si $a < V$ (caso 1) $H(Y) = h(p) + \log 2a$, lo que corresponde a la información de si se envió V o $-V$ más la entropía de Z.

   Si $a > V$ (caso 2) $H(Y) = \frac{V}{a}h(p) + \log 2a$
3. Información mutua $I = H(Y) - H(Y|X)$

   caso 1: $I = h(p)$. Capacidad $C = 1$ cuando son equiprobables.

   caso 2: $I = \frac{V}{a}h(p)$. Capacidad $C = \frac{V}{a}$. Decrece si a aumenta.

4. Umbral de decisión:

   caso 1: El error es 0 si $U \in (-V + a, V - a)$

   caso 2: El error es mínimo si U se elige de modo que se cometan menos errores en el símbolo más probable.

$$P_e = \frac{1-p}{2a}(-V + a - U) + \frac{p}{2a}(U - (V - a))$$

   Si $1 - p > p$, o sea $p < \frac{1}{2}$ el umbral óptimo es $U = -(V - a)$

   Si $p > \frac{1}{2}$, umbral óptimo $U = V - a$

Solución problema 3

1. a) canal sin errores; b, c) canal Z

2. a) canal sin errores; b, c) canal Z con memoria

3. Es la capacidad del canal Z sin memoria. Se resuelve igualando a cero $I(X;Y)$ y resolviendo para $P(X = 1)$. El resultado es $C = H(1/5) - 2/5 \approx 0{,}322$. La dist. que lo maximiza es $P^*(X = 1) = 2/5$

4. En b), si ocurre un error en tiempo i, entonces no puede haber error en el tiempo $i + 1$. Se puede entonces usar las muestras en los tiempos $i + 1$ como un canal sin errores y aumentar la capacidad del canal significativamente. En el caso c), si ocurre un error en el tiempo i, puede ocurrir un error en el tiempo $i + 1$ pero no en el $i + 2$. Procediendo de manera similar se puede aumentar la capacidad con respecto al modelo sin memoria.