Comunicación de datos
Curso 2016/17, Problemas # 1

1. (Propiedad de partición) Calcule la entropía de las siguientes distribuciones de probabilidad:

a) $\vec{p} = (1/3, 1/3, 1/9, 1/9, 1/27, 1/27, 1/27)$

b) $\mathbb{P}(X = k) = (1-\rho)\rho^k$ para $k \geq 0$, $0 < \rho < 1$. Use la igualdad $\sum_{k=0}^{\infty} k\rho^k = \frac{\rho}{(1-\rho)^2}$ para $|\rho| < 1$.

c) $\mathbb{P}(X = k) = A\rho^k$ para $0 \leq k \leq n$. Puede aplicar el resultado del apartado anterior y la propiedad de partición. Distinga los casos $0 < \rho < 1$ y $\rho \geq 1$.

d) $\vec{p} = (1-a, ap_1, \ldots, ap_n)$, en donde $(p_1, \ldots, p_n)$ es una distribución de probabilidad arbitraria y $a \in (0, 1)$. ¿Para qué valor de a es máxima la entropía?

2. (Canales y entropía)

a) Considere un canal BSC(p) cuyos símbolos de entrada X tienen probabilidades α y $1-\alpha$. Si Y denota la salida del canal, utilice la convexidad de la función de entropía para mostrar que $H(Y) \geq H(X)$. En consecuencia, si la entrada es uniforme, la salida también lo será.

b) Calcule la entropía de la salida de un canal BEC(ε) cuyas entradas tienen pobabilidades α y $1-\alpha$. ¿Es mayor o menor que la entropía de la entrada?

3. Sea X una variable aleatoria discreta *con un rango arbitrario*. ¿Qué entropía es mayor?

a) $H(X)$ o $H(10X)$

b) $H(X)$ o $H(X^2)$

c) $H(X^2)$ o $H(X^3)$

d) $H(X)$ o $H(X^+, X^-)$, siendo $X^+ = \big(|X| + X\big)/2$ y $X^- = \big(|X| - X\big)/2$.

4. (Desigualdades). Si la desigualdad es verdadera, demuéstrela; si no lo es, dé un contraejemplo.

a) $H(X, Y \mid Z) \geq H(X \mid Z)$.

b) $H(X \mid Z) \leq H(Z)$.

c) $H(X \mid Z) \leq H(X) - H(Z)$.

d) $H(X_1 \mid X_1 + X_2 + X_3) \geq H(X_1 + X_2 \mid X_1 + X_2 + X_3)$ si X_1, X_2 y X_3 son variables *iid*.

e) **(La entropía de las sumas parciales)** Si $X_1, \ldots, X_n$ son variables *iid*, $H(X_1 \mid X_1 + \cdots + X_n) \geq H(X_1 + X_2 \mid X_1 + \cdots + X_n) \geq \cdots \geq H(X_1 + \cdots + X_n \mid X_1 + \cdots + X_n) = 0$.

5. Por un canal discreto sin memoria con matriz de probabilidades de transición

$$Q = \frac{1}{6}\begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

se transmiten símbolos de acuerdo con la distribución de probabilidad $(1/2, 1/4, 1/4)$. Si X denota la entrada al canal e Y su salida, calcule:

a) $H(X)$ y $H(Y)$

b) $H(Y \mid X)$ y $H(X, Y)$

c) $H(X, Y)$ si se sabe que no se transmite x_1 y no se recibe y_1. **Nota**: la fuente y el canal siguen siendo los mismos.

6. Considere la matriz de probabilidades conjuntas del par de variables (X, Y)

$$\frac{1}{50}\begin{pmatrix} 2 & 6 & 1 & 3 \\ 5 & 2 & 8 & 0 \\ 1 & 7 & 3 & 4 \\ 0 & 4 & 4 & 0 \end{pmatrix}$$

en donde X indexa las filas e Y las columnas.

a) Calcule $H(X, Y)$, $H(X \mid Y)$, $H(Y \mid X)$.

b) Suponga que se nos dice que X es par y que $Y \neq 1$, y sean (X', Y') las variables con los valores ahora posibles de X e Y. Calcule $H(X', Y')$, $H(X' \mid Y')$, $H(Y' \mid X')$. **Nota**: suponga que las probabilidades conjuntas de los pares considerados mantienen las proporciones.

7. Considere el canal de matriz de probabilidades de transición

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1-p & p & 0 & 0 \\ 0 & 0 & p & 1-p & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

a) Sean X e Y las variables aleatorias de entrada y salida del canal, respectivamente, y sea $\mathbf{p} = (1/4, 1/4-\epsilon/2, \epsilon, 1/4-\epsilon/2, 1/4)$ el vector de probabilidades de Y. ¿Cuánto ha de valer p para que $I(X; Y) = 7/4$ bits?

b) ¿Es posible obtener a la salida símbolos equiprobables? Si lo fuese, calcule la información mutua para ese caso.

8. Considere el canal discreto sin memoria con matriz de probabilidades de transición

$$\begin{pmatrix} 1 & 0 & 0 \\ p/2 & 1-p & p/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

X e Y designan, respectivamente, la entrada y la salida de este canal.

a) Suponga que X es uniforme. Calcule $H(X, Y)$ y determine el valor de p que la maximiza.

b) Suponga que $p = 1$ y que los símbolos de X que se transmiten fielmente son equiprobables. Calcule el máximo de $I(X; Y)$ y determine para qué distribución de X se alcanza.

9. Una urna contiene r bolas rojas y b blancas. Teniendo en cuenta el orden de extracción de las bolas,

a) ¿Qué experimento tiene mayor entropía, extraer dos bolas con o sin restitución? Conteste sin realizar el cálculo de las entropías.

b) Calcule ahora las entropías de ambos experimentos.

c) Suponga que se extraen, consecutivamente y sin restitución, dos bolas de dicha urna. Determine la cantidad de información de la segunda de las extracciones si no se conoce el color de la bola extraída en primer lugar.

10. Una competición entre dos equipos queda resuelta cuando uno de ellos gana cuatro partidos al otro. Sea X la variable aleatoria que representa los resultados de los partidos de una competición entre los equipos A y B; valores posibles de X son AAAA, BABABAB, etc. Sea Y el número de partidos jugados, que varía entre 4 y 7. Suponiendo que ambos equipos tienen la misma probabilidad de ganar un partido y que los partidos son independientes, calcule $H(X), H(Y), H(Y \mid X)$ y $H(X \mid Y)$.

11. Los niños de un colegio pertenecen a uno y sólo uno de los equipos que se han formado para practicar diversos deportes. Sea n el número de equipos, y sean la mitad de ellos de fútbol (11 jugadores), la cuarta parte de balonmano (7 jugadores) y el resto de baloncesto (5 jugadores). Suponga que todos los equipos tienen nombres diferentes y que los jugadores de un mismo equipo se identifican por el número de su camiseta (números naturales consecutivos empezando por 1). Elija uno de los niños del colegio al azar y averigüe su deporte, equipo y número de camiseta.

a) ¿Cuánta información proporciona el deporte?

b) ¿Cuánta información proporcionan el deporte y el equipo conjuntamente?

c) ¿Cuánta información proporcionaría el equipo si el deporte fuese conocido?

d) ¿Aporta alguna información el deporte sobre el número de la camiseta?

e) ¿Cuál es la incertidumbre del deporte después de conocer el número de la camiseta?

12. Un estudio sobre el personal de cierta empresa de telecomunicaciones revela que el 70 % de sus ingenieros son hombres, y que el 25 % de la plantilla no rebasa los 30 años de edad (≤ 30), con un 20 % mayor de 40. Además, el número de mujeres en este último tramo (> 40) es despreciable, y las más jóvenes (≤ 30) duplican al resto de sus compañeras.

a) Considerando únicamente esta información, ¿qué incertidumbre existe a priori sobre el grupo humano (definido por sexo e intervalo de edad) donde se situará una nueva contratación?

b) ¿Y si nos dicen que la nueva contratación no es un hombre mayor de 30 años?

13. De un grupo de estudiantes, el 25 % no están preparados para la universidad. Sin embargo, el 25 % de estos son aceptados tras una prueba de selección, que es superada por el 50 % de los estudiantes.

a) Al oír el resultado de la prueba, ¿cuánta información recibe un estudiante no preparado?

b) Ídem si la selección se decide tirando una moneda al aire.

14. Los alumnos de un curso pueden dividirse, en función de su capacidad, en dos grupos de igual tamaño. Los del grupo A resuelven satisfactoriamente el 75 % de las cuestiones que se les plantean; mientras que los del grupo B sólo resuelven el 50 %. Suponga que se elige al azar un alumno del curso y se le somete a un examen de tres preguntas independientes.

a) ¿Qué aporta más información para clasificar a un alumno, que responda todas bien o que no responda bien ninguna?

b) ¿Cuánta información proporciona la respuesta a una sola pregunta?

15. Los habitantes de un pueblo se dividen en dos grupos, A y B. La mitad de las personas del grupo A siempre dicen la verdad, $3/10$ mienten siempre y las demás siempre rehúsan contestar. En el grupo B, $3/10$ de las personas son sinceras, la mitad son mentirosas y las demás siempre rehúsan contestar. Sea p la probabilidad de que una persona elegida al azar sea del grupo A. Sea $I = I(p)$ la información que sobre el tipo de respuestas de una persona proporciona su pertenencia a un grupo. Determine el valor máximo posible de I y el porcentaje de personas en el grupo A para el que dicho máximo se da.