Compresión de datos sin pérdida
Práctico 1

Ejercicio 1
Consideramos un alfabeto de tres símbolos, $\mathcal{A} = \{a, b, c\}$, y dos distribuciones p, q, definidas según la siguiente tabla.

Símbolo	$p(x)$	$q(x)$
a	$\frac{1}{2}$	$\frac{1}{3}$
b	$\frac{1}{4}$	$\frac{1}{3}$
c	$\frac{1}{4}$	$\frac{1}{3}$

Calcular $H(p)$, $H(q)$, $D(p||q)$ y $D(q||p)$. Observar que $D(p||q) \neq D(q||p)$; dar un ejemplo de distribuciones distintas sobre un alfabeto binario tales que se de la igualdad.

Ejercicio 2 *Entropía de funciones de variables aleatorias*
Sea X una variable aleatoria discreta que toma valores en $\mathcal{A} = \{a_1, a_2, \ldots a_m\}$, con distribución dada por el vector de probabilidad $p = (p_1, p_2, \ldots, p_m)$. Sea $g : \mathcal{A} \to \mathcal{B}$ una función.

1. Mostrar que $H(X) \geq 0$ y determinar todos los valores de p tales que $H(X) = 0$.
2. Probar que si Y es una variable aleatoria sobre un alfabeto $\mathcal{Y}$ tal que $H(Y|X) = 0$, entonces Y es función de X, es decir, para todo $x \in \mathcal{A}$ tal que $p(x) > 0$, existe un único $y \in \mathcal{Y}$ que cumple $p(x, y) > 0$.
3. Descomponer $H(X, g(X))$, usando la regla de la cadena de dos formas distintas, y concluir que $H(g(X)) \leq H(X)$, con igualdad si y sólo si g restringida al subconjunto con probabilidad positiva de $\mathcal{A}$ es inyectiva.

Ejercicio 3 *Máxima entropía sobre alfabetos finitos*
Sea X una variable aleatoria sobre un alfabeto finito $\mathcal{A}$ distribuida según p. Probar que $H(X) \leq \log |\mathcal{A}|$ con igualdad si y sólo si $p = u$, donde u es la distribución uniforme sobre $\mathcal{A}$.
Sugerencia: Considerar la divergencia entre p y u y usar la desigualdad de la información.

Ejercicio 4 *Cota de independencia*
Sean X, Y variables aleatorias discretas. Probar que $H(X, Y) \leq H(X) + H(Y)$ con igualdad si y sólo si X, Y son independientes.
Sugerencia: Usar la desigualdad de la información.

Ejercicio 5 *Condicionar no aumenta la entropía*
Sean X, Y variables aleatorias discretas. Probar que $H(X|Y) \leq H(X)$ con igualdad si y sólo si X, Y son independientes.
Sugerencia: Usar la cota de independencia y la regla de la cadena.

Ejercicio 6 *Máxima entropía sobre* $\mathbb{N}$ *para esperanza dada*
Sea X una variable aleatoria sobre el alfabeto de los números naturales, $\mathbb{N}$, con valor esperado $E[X] = \mu$.

1. Probar que, entre todas las distribuciones p sobre $\mathbb{N}$ que satisfacen $E_p[X] = \mu$, la entropía $H(X)$ se maximiza cuando p es de la forma $p(n) = 2^{\lambda_0 + \lambda_1 n}$.
   Sugerencia: Considerar una ditribución arbitraria q sobre $\mathbb{N}$ tal que $E_q[X] = \mu$, usar que $D(q||p) \geq 0$ y el hecho de que $E_q[X] = E_p[X] = \mu$.

2. Despejar λ_0 y λ_1 en función de μ, usando que $\sum_{n\in\mathbb{N}} p(n) = 1$ y $E_p[X] = \mu$.
   Sugerencia: Recordar que

$$\sum_{n\in\mathbb{N}} r^n = \frac{1}{1-r}\,, \quad \sum_{n\in\mathbb{N}} nr^n = \frac{r}{(1-r)^2}\,.$$

3. Concluir que $H(X) \leq (1+\mu)\log(1+\mu) - \mu\log\mu\,.$

Ejercicio 7
Sea $\mathcal{A} = \{a_1, a_2, \dots a_m\}$ un alfabeto finito y Q una distribución de probabilidad sobre $\mathcal{A}^n$ dada por

$$Q(x^n) = \prod_{i=1}^{n} q(x_i)\,,$$

donde q es una distribución sobre $\mathcal{A}$ (los símbolos de la secuencia se generan i.i.d. según q). Sea $\hat{p}(x^n)$ la distribución de probabilidad empírica sobre $\mathcal{A}$ inducida por la secuencia x^n, esto es,

$$\hat{p}(x^n)(a) = \frac{1}{n}|\{i = 1 \dots n : x_i = a\}|\,, \quad a \in \mathcal{A}\,.$$

Mostrar que $Q(x^n)$ se puede escribir como

$$Q(x^n) = 2^{-n\left[H(\hat{p}(x^n)) + D(\hat{p}(x^n)||q)\right]}\,.$$

Ejercicio 8 *Run length coding*
Sean $X_1, X_2, \dots, X_n$ variables aleatorias binarias (no necesariamente independientes). Supongamos que para esta secuencia calculamos los largos de corridas (ocurrencias consecutivas de un mismo símbolo), $R = (R_1, R_2, \dots)$, en orden de ocurrencia. Por ejemplo, para la secuencia $X = 0001100100$, obtenemos los largos de corridas $R = (3, 2, 2, 1, 2)$. Compare $H(X_1, X_2, ..., X_n)$, $H(R)$, y $H(X_n, R)$. Muestre todas las igualdades y desigualdades y acote la diferencia en caso de desigualdad.