

UNIVERSIDAD NACIONAL DE SAN JUAN

FACULTAD DE CIENCIAS EXACTAS FÍSICAS Y NATURALES

DEPARTAMENTO DE INFORMÁTICA

LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN

Asignatura: **“Teorías de la Información”**

**Práctico 1**

**ALUMNOS:**

* **Juan Peña 21.151**
* **Nicolás Quiroga Santini 21.444**
* **Baltasar Ortiz Becerra 21.752**

**Equipo de Cátedra**

**Raul Klenzi**

**Manuel Ortega**

**Fabricio Amaya**

**AÑO 2025**

Teoría de Información - Práctico 1

## Ejercicio 1:

**Supón una imagen de resolución 2560 x 1440 píxeles, codificada en 30 bits por píxel. Calcula la cantidad de información que proporciona un cuadro de imagen. Luego, considera una imagen de 7680 x 4320 píxeles (resolución 8K) codificada en HDR con 12 bits por canal de color, ¿cuánta información se genera?**

Imagen 1: 2560 × 1440 píxeles, 30 bits/píxel

* **Píxeles totales**: 2560 × 1440 = 3,686,400 píxeles
* **Información total**: 3,686,400 × 30 = 110,592,000 bits
* **En bytes**: 13,824,000 bytes = **13.8 MB**

Imagen 2: 7680 × 4320 píxeles (8K HDR), 12 bits/canal × 3 canales

* **Píxeles totales**: 7680 × 4320 = 33,177,600 píxeles
* **Bits por píxel**: 12 × 3 = 36 bits/píxel
* **Información total**: 33,177,600 × 36 = 1,194,393,600 bits
* **En bytes**: 149,299,200 bytes = **149.29 MB**

## Ejercicio 2:

**Un narrador utiliza 1000 palabras tomadas al azar de un vocabulario de 20.000 palabras. Calcula la información generada. Luego, compárala con la información contenida en una única imagen de resolución 640x480 píxeles codificada en 24 bits por píxel (color verdadero). Analiza cuál contiene más información y verifica si se cumple el dicho popular 'una imagen vale más que mil palabras' incluso con baja resolución..**

Narrador: 1000 palabras de vocabulario de 20,000

* **Información por palabra**: log₂(20,000) = 14.29 bits
* **Información total**: 1000 × 14.29 = **14,287.71 bits**

Imagen 640×480, 24 bits/píxel

* **Píxeles**: 640 × 480 = 307,200
* **Información total**: 307,200 × 24 = **7,372,800 bits**

Conclusión

La imagen contiene **516 veces más información** que las mil palabras, confirmando que **"una imagen vale más que mil palabras"**

## Ejercicio 3:

**Calcula la información generada por un mensaje de 200 caracteres tomados al azar de un alfabeto de 32 símbolos. ¿Y por un mensaje de 200 caracteres con un alfabeto de 64 símbolos? Ahora repite el cálculo para un mensaje de 400 caracteres con alfabetos de 32 y 64 símbolos. Analiza cómo cambia la cantidad de información y verifica si duplicar el tamaño del alfabeto duplica la información total..**

Mensajes de 200 caracteres:

* **Alfabeto 32 símbolos**: log₂(32) = 5 bits/carácter → **1000 bits total**
* **Alfabeto 64 símbolos**: log₂(64) = 6 bits/carácter → **1200 bits total**

Mensajes de 400 caracteres:

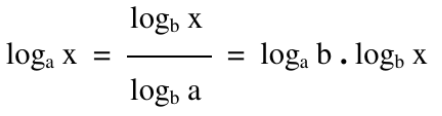
* **Alfabeto 32 símbolos**: **2000 bits total**
* **Alfabeto 64 símbolos**: **2400 bits total**

Análisis

* **Relación 64/32 símbolos**: 1.2 (constante)
* **Duplicar el alfabeto NO duplica la información total**
* La relación es: log₂(64)/log₂(32) = 6/5 = 1.2

## Ejercicio 4:

**Demostrar la siguiente igualdad**:



**Demostración:**

**Paso 1:** Definiendo y = log\_a x

Por definición de logaritmo: **a^y = x**

**Paso 2:** Aplicando logaritmo en base b a ambos lados

log\_b(a^y) = log\_b(x)

**Paso 3:** Usemos la propiedad log\_b(m^n) = n·log\_b(m)

y · log\_b(a) = log\_b(x)

**Paso 4:** Despejando y

y = log\_b(x)/log\_b(a)

**Paso 5:** Luego sustituyendo y = log\_a x

**log\_a x = (log\_b x)/(log\_b a)**

**Para demostrar que log\_a x = log\_a b · log\_b x:**

Multiplicando ambos lados de nuestra fórmula por log\_b a:

log\_a x · log\_b a = log\_b x

Dividiendo por log\_b a:

log\_a x = log\_b x / log\_b a

Pero también sabemos que 1/log\_b a = log\_a b (propiedad de logaritmos recíprocos)

Por tanto: **log\_a x = log\_a b · log\_b x**

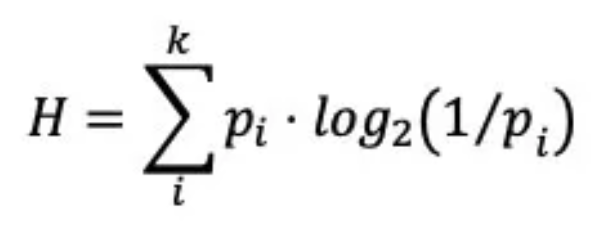
## Ejercicio 5:

**Una fuente F tiene 6 símbolos con probabilidades: p1=0.4, p2=0.2, p3=0.15, p4=0.1, p5=0.1, p6=0.05. Calcula la información individual y la entropía de la fuente.**

Fuente F con probabilidades: [0.4, 0.2, 0.15, 0.1, 0.1, 0.05]

**Información individual I(sᵢ) = log₂(1/p(sᵢ)):**

* I(s₁) = log₂(1/0.4) = 1.322 bits
* I(s₂) = log₂(1/0.2) = 2.322 bits
* I(s₃) = log₂(1/0.15) = 2.737 bits
* I(s₄) = I(s₅) = log₂(1/0.1) = 3.322 bits
* I(s₆) = log₂(1/0.05) = 4.322 bits



**Entropía de la fuente**: H(F) = **2.2842 bits**

## Ejercicio 6:

**Justifica por qué una distribución uniforme maximiza la entropía. Compara un dado justo (6 caras, p=1/6) con un dado sesgado: p1=0.3, p2=0.25, p3=0.15, p4=0.15, p5=0.1, p6=0.05**

La distribución uniforme maximiza la entropía porque:

La entropía mide la incertidumbre. La incertidumbre es máxima cuando todos los eventos son igualmente probables.

Comparación:

* **Dado justo** (p = 1/6 para todas las caras): H = **2.5850 bits**
* **Dado sesgado** [0.3, 0.25, 0.15, 0.15, 0.1, 0.05]: H = **2.3905 bits**

**Diferencia**: 0.1945 bits (el dado justo tiene mayor entropía)

## Ejercicio 7

**Sean 12 monedas una de las cuales tiene peso diferente, indicar cuántas pesadas son necesarias para encontrarla, especifique la metodología de peso**

**Estados posibles**: 12 monedas × 2 estados = 24 estados  
**Resultados por pesada**: 3 (>, =, <)  
**Pesadas mínimas**: ⌈log₃(24)⌉ = **3 pesadas**

1. **Primera pesada**: {1,2,3,4} vs {5,6,7,8}, fuera {9,10,11,12}
2. **Según resultado**, dividir en subcasos
3. **Tercera pesada** determina la moneda exacta y si es pesada/liviana

**Respuesta**: Se necesitan exactamente **3 pesadas**.

## Ejercicio 8

**Calcula la información de una letra al azar de un alfabeto de 27 símbolos (incluyendo ñ). Luego, de pares y ternas de letras.**

**Alfabeto español (27 símbolos: 26 letras + ñ):**

* **Una letra**: log₂(27) = **4.755 bits**
* **Par de letras**: 2 × 4.755 = **9.510 bits**
* **Terna de letras**: 3 × 4.755 = **14.265 bits**

**Estados posibles:**

* **Una letra**: 27 estados
* **Par de letras**: 27² = 729 estados
* **Terna de letras**: 27³ = 19,683 estados

*Estos cálculos asumen independencia e equiprobabilidad.*

## Ejercicio 9

**¿Cómo serán las cifras anteriores respecto de la forma castellana de escritura cotidiana? Si a su criterio existiera diferencia ¿a qué se debería?**

**Diferencias esperadas:**

1. **Frecuencias no uniformes**: Vocales (~45%), consonantes comunes (~25%), raras (<1%)
2. **Entropía real**: ~4.0 bits/letra vs 4.807 teórica (reducción del 16.8%)

**Causas:**

* Restricciones fonológicas
* Patrones morfológicos
* Frecuencia de palabras comunes
* Redundancia natural

## Ejercicio 10

**Explica las propiedades de la cantidad de información. Da un ejemplo para cada una de estas propiedades.**

**I(x) = log₂(1/p(x)) tiene estas propiedades:**

1. **No negatividad**: I(x) ≥ 0
2. **Monotonía decreciente**: Si p(x₁) > p(x₂), entonces I(x₁) < I(x₂)
3. **Aditividad**: I(x,y) = I(x) + I(y) para eventos independientes
4. **Máxima para equiprobables**: I = log₂(n)
5. **Límites**: Mínimo 0 (p=1), máximo ∞ (p→0)

## Ejercicio 11

**Dada una variedad V = 2000 sucesos encontrar la base b óptima de representación que hace mínima la siguiente relación número más corto, es decir que minimice I . b (cantidad de información por base).**

**Análisis matemático**:

f(b) = [log(V)/log(b)] · b

f'(b) = 0 → log(b) = 1 → **b = e ≈ 2.718**

**Comparación**:

* Base e: I·b = 20.661
* **Base 3**: I·b = 20.756 (mejor entero)
* Base 2: I·b = 21.932

Base óptima teórica es **e**, práctica es **base 3**.

## Ejercicio 12

**Para el texto contenido en el código QR de la parte superior de la página encuentre la entropía de orden cero o independiente tanto de lo relevado en el QR como en el texto correcto. ¿Qué conclusiones saca?**

**Texto:**

Sgeun un etsduio de una uivenrsdiad ignlsea, no ipmotra el odren en el que las ltears etsan ersciats, la uicna csoa ipormtnate es que la pmrirea y la utlima ltera esten ecsritas en la psiocion cocrrtea. El rsteo peuden estar ttaolmntee mal y aun pordas lerelo sin pobrleams. Etso es pquore no lemeos cada ltera por si msima preo la paalbra es un tdoo. Pesornamelnte me preace icrneilbe…  
  
Según un estudio de una universidad inglesa, no importa el orden en el que las letras están escritas, la única cosa importante es que la primera y la última letra estén escritas en la posición correcta. El resto pueden estar totalmente mal y aun podrás leerlo sin problemas. Esto es porque no leemos cada letra por si misma pero la palabra es un todo. Personalmente me parece increíble...

**La resolución se puede resolver con el código del práctico de máquina en el ejercicio número 6, reemplazando el contenido del archivo .txt con cada texto individualmente. Enlace directo del código que se puede utilizar:**

<https://github.com/BaltasarOrtiz/GRUPO_2_TDI_2025/blob/main/Practico_1/P1_M%C3%A1quina/ejercicio_3/ej_3.py>

**Resultados aproximados:**

* **Con errores: 3.9000 bits/carácter**
* **Sin errores: 4.0153 bits/carácter**

**Conclusión:** podemos decir que las entropías son muy similares, indicando que mezclar letras mantiene la distribución estadística del idioma español.