Teoría de la información

Entropía diferencial

Vanessa Martínez Romero, Moisés Emiliano Arellano Ávila y Fernando Gómez Perera

19/10/2020

Entropía diferencial

La entropía de Shannon como se mencionó en los artículos que se revisaron al inicio del curso se define como el valor esperado de $-\log(p_X(k))$, al cual se le llama sorpresa. Por lo tanto la entropía de Shannon para la variable aleatoria X queda especificado por la siguiente fórmula:

$$H(X) = -\sum_k p_X(k)\log(p_X(k)).$$

Nótese que esta fórmula está definida para distribuciones discretas y por lo tanto para el caso de las variables aleatorias continuas se tiene que utilizar otra definición. Shannon, extendió la entropía definida arriba mediante una simple extensión de sumatoria a integral y obtuvo lo que se llama entropía diferencial de X, la cual está dada por:

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

Esta entropía, en principio, se utilizó para el cálculo de la entropía para distribuciones continuas aunque ahora existen alternativas más precisas para el cálculo de la entropía de distribuciones continuas.

Entropía diferencia de distribución uniforme

Basándonos en lo anterior, calcularemos la entropía diferencial para algunas distribuciones conocidas y verificaremos el efecto que tienen ciertos parámetros en la entropía diferencial. Recordemos que la distribución uniforme tiene la siguiente fórmula:

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{en otro caso} \end{cases}$$

El gráfico que define a la distribución uniforme $\mathcal{U}(2,4)$ es, por ejemplo, el siguiente:

```
t <- seq(1, 5, length = 200)
original <- ifelse(t >= 2 & t <= 4, 1, 0)
highchart() %>% hc_add_series(cbind(t, original), name = "Densidad uniforme") %>%
    hc_add_theme(hc_theme_smpl()) %>% hc_title(text = "f(x) = 1/(b-a)") %>% hc_subtitle(text = "IT0322 
    hc_xAxis(title = list(text = "Tiempo")) %>% hc_yAxis(title = list(text = "Valores de f(x)"))
```

La entropía diferencial de $\mathcal{U}(a,b)$ es por lo tanto:

$$
\begin{aligned}
h(X) &= -\int_a^b \left(\frac{1}{b-a}\right)\log(\frac{1}{b-a})\,dx \\
&= \frac{\log(b-a)}{b-a}\int_a^b dx \\
&= \frac{\log(b-a)}{b-a} \times (b-a) \\
h(X) &= \log(b-a)
\end{aligned}
$$

Es decir, la entropía diferencial de la distribución uniforme es:

$$h(X) = \log(b-a)$$

Ahora sabemos que la varianza de la distribución uniforme es:

$$Var(X) = \frac{(b-a)^2}{12}$$

entonces, podemos relacionar la varianza de la distribución con la entropía. Supongamos que $a = 0$ y b es variables, entonces, el gráfico de la varianza y la entropía son:

```
b <- seq(1, 6, length = 100)
a <- 0
varX <- (b - a)^2/12
hX <- log(b - a)
highchart() %>% hc_add_series(cbind(b, hX), name = "Entropía diferencial") %>% hc_add_series(cbind(b,
    varX), name = "Varianza") %>% hc_add_theme(hc_theme_smpl()) %>% hc_title(text = "Entropía y varianza
    hc_subtitle(text = "IT0322 - Teoría de la información") %>% hc_xAxis(title = list(text = "Tiempo"))
    hc_yAxis(title = list(text = "Valores"))
```

De lo anterior se puede concluir que la entropía incrementa con la varianza de la distribución uniforme, es decir mientras $b - a$ incremente también lo hará la entropía.

Ejercicios

1. Suponga ahora que la varianza es constante pero la distribución se traslada sobre diversos puntos de x. ¿Cómo es el comportamiento de la entropía en este caso?. Sugerencia: grafique la traslación contra entropía para obtener la respuesta.

Respuesta:

Para que la varianza sea constante, la diferencia entre a y b debe mantenerse constante. Para ello, se define a con valores de -5 a 5, y b como $b = a + 2$.

La traslación x se define como $x = \frac{a+b}{2}$, que es el punto intermedio entre a y b.

```
a <- seq(-5, 5, length = 100)
b <- a + 2
x <- (a + b)/2
varX <- (b - a)^2/12
hX <- log(b - a)
highchart() %>% hc_add_series(cbind(x, hX), name = "Entropía diferencial") %>% hc_add_series(cbind(x,
    varX), name = "Varianza") %>% hc_add_theme(hc_theme_smpl()) %>% hc_title(text = "Ejercicio 1: Entrop
    hc_subtitle(text = "IT0322 - Teoría de la información") %>% hc_xAxis(title = list(text = "Traslación
    hc_yAxis(title = list(text = "Valores"))
```

Si la varianza se mantiene constante, entonces la entropía también se mantiene constante.

2. Suponga que la densidad de una variable aleatoria X está dada por:

$$f(x) = \begin{cases} \frac{7}{4} - \frac{3}{2}x & 0 < x < 1 \\ 0 & \text{otro caso} \end{cases}$$

Hallar $h(X)$.

Respuesta:

La entropía para la densidad de esta variable aleatoria se obtiene por medio de la siguiente integral:

$$h(X) = -\int_0^1 \left(\frac{7}{4} - \frac{3}{2}x\right) log\left(\frac{7}{4} - \frac{3}{2}x\right) dx$$

```
fx <- function(x) 7/4 - 3 * x/2
hX <- integrate(function(x) -fx(x) * log(fx(x)), lower = 0, upper = 1)
hX
```

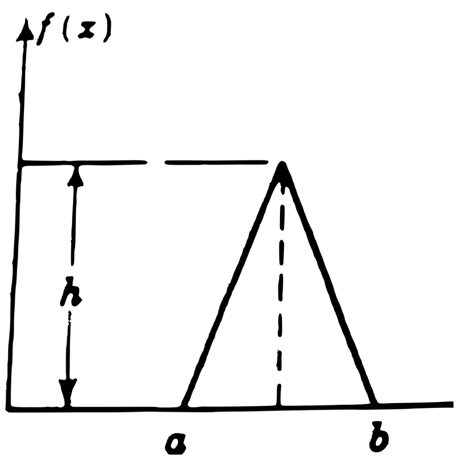
```
-0.1001556 with absolute error < 2.2e-11
```

Por lo tanto, se obtiene que

$$h(X) = -0.1002$$

Ejercicios

1. Hallar la entropía para la siguiente función de densidad:



Respuesta:

La función de densidad es:

$$f(x) = \begin{cases} \frac{h}{c-a}(x-a) & a \le x < c \\ h & x = c \\ \frac{h}{b-c}(b-x) & c < x \le b \\ 0 & \text{en otro caso} \end{cases}$$

donde

$$h = \frac{2(c-a)}{(a-c)^2 + (c-a)(b-c)}$$

y

$$c = \frac{a+b}{2}$$

Al sustituir c en h, y simplificando este resultado, se obtiene que

$$h = \frac{2}{b-a}$$

quedando la función de densidad como

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & a \leq x < c \\ \frac{2}{b-a} & x = c \\ \frac{2(b-x)}{(b-a)(b-c)} & c < x \leq b \\ 0 & \text{en otro caso} \end{cases}$$

que corresponde a la función de densidad de la distribución triangular.

```
# Funciones para C
C <- function(A, B) (A + B)/2

# Funcion de densidad
fx <- function(A, B, x) {
    C <- C(A, B)
    if (A <= x & x < C) {
        return(2 * (x - A)/(B - A) * (C - A))
    } else if (x == C) {
        return(2/(B - A))
    } else if (C < x & x <= B) {
        return(2 * (B - x)/(B - A) * (B - C))
    } else {
        return(0)
    }
}
```

Para saber si el cálculo de esta función de densidad es correcto, se toma como ejemplo $a = 3$ y $b = 5$, y se calcula si $\int_{-\infty}^{+\infty} fx(x)dx = 1$.

```
# Funcion de prueba
ex_fx <- function(x) {
    a <- 3
    b <- 5
    res <- c()
    for (i in x) {
        res <- append(res, fx(a, b, i))
    }
    return(res)
}

integrate(ex_fx, -Inf, Inf)
```

```
1 with absolute error < 2.7e-06
```

Este ejemplo genera la siguiente gráfica:

```
t <- seq(0, 6, length = 200)
original <- ex_fx(t)
highchart() %>% hc_add_series(cbind(t, original), name = "Densidad") %>% hc_add_theme(hc_theme_smpl()) %
    hc_title(text = "Función de densidad con a = 3 y b = 5") %>% hc_subtitle(text = "IT0322 - Teoría de
    hc_xAxis(title = list(text = "Tiempo")) %>% hc_yAxis(title = list(text = "Valores de f(x)"))
```

Con la función comprobada, finalmente se calcula la entropía diferencial de esta distribución:

$$h(X) = -\int_a^c \left(\frac{2(x-a)}{(b-a)(c-a)}\right) log\left(\frac{2(x-a)}{(b-a)(c-a)}\right)dx$$

$$-\int_c^b \left(\frac{2(b-x)}{(b-a)(b-c)}\right) log\left(\frac{2(b-x)}{(b-a)(b-c)}\right)dx$$

donde el resultado de ambas integrales queda como

$$h(X) = -\frac{\left(2log\left(\frac{2}{b-a}\right)-1\right)(c-a)}{2(b-a)} - \frac{\left(2log\left(\frac{2}{b-a}\right)-1\right)(b-c)}{2(b-a)}$$

que finalmente se reduce a

$$h(X) = \frac{1}{2} + log\left(\frac{b-a}{2}\right)$$

2.- Para la densidad anterior, ¿cuál es la relación entre la varianza y la entropía? ¿Existe alguna relación entre la altura h y la entropía $h(X)$?

Respuesta:

Para obtener la varianza, primero es necesario obtener la media (que es el primer momento de la función) usando la fórmula $\mu = \int_{-\infty}^{\infty} xf(x)dx$.

$$\mu = \int_a^c x\left(\frac{2(x-a)}{(b-a)(c-a)}\right)dx + \int_c^b x\left(\frac{2(b-x)}{(b-a)(b-c)}\right)dx$$

donde el resultado de las integrales queda como

$$\mu = \frac{2c^3-2a^3-3ac^2+3a^3}{3(b-a)(c-a)} + \frac{b^3-3bc^2+2c^3}{3(b-a)(b-c)}$$

que finalmente se reduce a

$$\mu = \frac{a+b+c}{3}$$

Con este resultado, se puede obtener la varianza (que es el segundo momento de la función) por medio de la fórmula $\sigma^2 = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$.

$$\sigma^2 = \int_a^c \left(x - \frac{a+b+c}{3}\right)^2 \left(\frac{2(x-a)}{(b-a)(c-a)}\right)dx$$

$$+\int_c^b \left(x - \frac{a+b+c}{3}\right)^2 \left(\frac{2(b-x)}{(b-a)(b-c)}\right) dx$$

donde el resultado de las integrales queda como

$$\sigma^2 = \frac{(c-a)(3c^2 - 2c(2b+a) + 2b^2 + a^2)}{18(b-a)}$$

$$-\frac{(c-b)(3c^2 - 2c(b+2a) + b^2 + 2a^2)}{18(b-a)}$$

que finalmente se reduce a

$$\sigma^2 = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}$$

Gráficamente la relación entre la varianza y la entropía se visualiza de la siguiente forma:

```
b <- seq(1, 10, length = 200)
a <- 0
c <- C(a, b)
hX <- 1/2 + log((b - a)/2)
varX <- (a^2 + b^2 + c^2 - a * b - a * c - b * c)/18
highchart() %>% hc_add_series(cbind(b, hX), name = "Entropía diferencial") %>% hc_add_series(cbind(b,
    varX), name = "Varianza") %>% hc_add_theme(hc_theme_smpl()) %>% hc_title(text = "Relación entre la V
    hc_subtitle(text = "Teoría de la información") %>% hc_xAxis(title = list(text = "Tiempo")) %>%
    hc_yAxis(title = list(text = "Valores"))
```

Con este gráfico se puede concluir que mientras la varianza aumenta, la entropía también aumenta. La entropía aumenta de forma logarítmica, mientras que la varianza aumenta de forma exponencial.

Gráficamente la relación entre la altura h y la entropía $h(X)$ se visualiza de la siguiente forma:

```
b <- seq(1, 10, length = 200)
a <- 0
h <- 2/(b - a)
hX <- 1/2 + log((b - a)/2)
highchart() %>% hc_add_series(cbind(h, hX), name = "Entropía") %>% hc_add_theme(hc_theme_smpl()) %>%
    hc_title(text = "Entropía vs. Altura") %>% hc_subtitle(text = "Teoría de la información") %>%
    hc_xAxis(title = list(text = "Altura")) %>% hc_yAxis(title = list(text = "Entropía"))
```

La altura sí modifica la entropía, y la relación entre ambas variables es que la entropía disminuye mientras la altura aumenta.