Igualdades en Análisis de LCFS hashing

Matías Salim Seda Auil
Universidad de Chile

El siguiente trabajo tiene como propósito demostrar la siguiente igualdad relacionada al análisis de LCFS hashing:

$$P_n(z) = \frac{\bar{B_n}(z) - \bar{B_n}(0)}{1 - \bar{B_n}(0)}$$

Para desarrollar lo pedido se utilizará una herramienta de álgebra computacional (Maple)[(1)] y se mostará que la igualdad es cierta siempre que $n \in \mathbb{N}$, $m > 0$ y $m \geq n$.

PRESENTACIÓN DEL PROBLEMA: ELEMENTOS DEL ANÁLISIS LCFS DE HASHING.

Se sabe que la función generatriz de probabilidad $R_n(x_1, ..., x_n)$ del incremento del costo de cada una de las llaves como resultado de una inserción de una K_n llave viene dado por la siguiente expresión:

$$R_n(x_1, ..., x_n) = (m - n + 1)\frac{\frac{x_n}{m+x_m}}{1 - \sum_{j=1}^{n} \frac{x_j}{x_j+m}}$$

Asimismo, se sabe que la función generatriz $Q_n(x_1, ..., x_n)$ del costo de búsqueda de cada llave producto de las inserciones K_1, K_2, ..., K_n (en ese orden) viene expresado por la siguiente igualdad:

$$Q_n(x_1, ..., x_n) = \prod_{k=1}^{n} R_k(x_1, ..., x_k)$$

Así, la función generatriz $P_n(z)$ del costo de búsqueda existoso es:

$$P_n(z) = \frac{1}{n}(Q_n(z, 1, ..., 1) + Q_n(1, z, ..., 1) + ... + Q_n(1, 1, ..., z))$$

Dada la información planteada, se quiere demostrar que:

$$P_n(z) = \frac{\bar{B_n}(z) - \bar{B_n}(0)}{1 - \bar{B_n}(0)}$$

donde:

$$\bar{B_n}(z) = \frac{\binom{m}{n}}{\binom{m+1-\bar{z}}{n}}$$

y

$$\bar{z} = z\frac{m+1}{m+z}$$

RESOLUCIÓN DEL PROBLEMA: RESOLVIENDO LAS DIFERENCIAS

Nótese que para R_k tal que z se encuentra en la i-ésima coordenada, con $1 \leq i \leq k-1$, y donde en las $k-1$ posiciones restantes hay un 1 se tiene que:

$$R_k(1, ..., z, ..., 1) = (m - k + 1)\frac{\frac{1}{m+1}}{1 - ((\sum_{j=1}^{k-1} \frac{1}{1+m}) + \frac{z}{z+m})}$$

Desarrollando la expresión se puede llegar a la siguiente igualdad:

$$R_k(1, ..., z, ..., 1) = \frac{(m+1-k)(m+z)}{m^2 - m(k-2) - z(k-1)}$$

Para el caso R_k tal que z se encuentra en la k-ésima coordenada y con 1's en las $k-1$ posiciones restantes, se tiene que:

$$R_k(1,1,...,z) = (m-k+1)\frac{\frac{z}{m+z}}{1-((\sum_{j=1}^{k-1}\frac{1}{1+m})+\frac{z}{z+m})}$$

Desarrollando la expresión se puede llegar a la siguiente igualdad:

$$R_k(1,1,...,z) = \frac{(m+1-k)(m+1)z}{m^2-m(k-2)-z(k-1)}$$

Asimismo, no es difícil notar que $R_k(1,1,...,1) = 1$.

Para Q_n tal que z se encuentra en la i-ésima coordenada, con $1 \le i \le n-1$ se tiene que:

$$Q_n(1,1,...,z,...,1) = (\prod_{k=1}^{i-1} R_k(1,1,...,z))R_i(1,1,...,z)(\prod_{k=i+1}^{n} R_k(1,1,...,z,...,1))$$

Lo que es equivalente a:

$$Q_n(1,1,...,z,...,1) = R_i(1,1,...,z)\prod_{k=i+1}^{n} R_k(1,...,z,...,1)$$

Donde i representa el *lugar* de la coordenada donde se encuentra z

Reemplazando $R_i(1,1,...,z)$ y $R_k(1,...,z,..,1)$ con los valores calculados anteriormente se tiene que:

$$Q_n(1,1,...,z,...,1) = \frac{(m+1-i)(m+1)z}{m^2-m(i-2)-z(i-1)}\prod_{k=i+1}^{n}\frac{(m+1-k)(m+z)}{m^2-m(k-2)-z(k-1)}$$

Desarrollando la expresión se cumple que:

$$Q_n(1,1,...,z,...,1) = (m+1)(m+z)z\prod_{k=i}^{n}\frac{m+1-k}{m^2-m(k-2)-z(k-1)}$$

Por lo tanto, se tiene que $P_n(z)$ viene dado por:

$$P_n(z) = \frac{1}{n}\sum_{i=1}^{n}(m+1)(m+z)z\prod_{k=i}^{n}\frac{m+1-k}{m^2-m(k-2)-z(k-1)}$$

Lo que es equivalente a:

$$P_n(z) = \frac{(m+1)(m+z)z}{n}\sum_{i=1}^{n}\prod_{k=i}^{n}\frac{m+1-k}{m^2-m(k-2)-z(k-1)}$$

Resolviendo la ecuación con un programa de álgebra computacional se obtiene lo siguiente:

$$P_n(z) = \frac{(z+m)\Gamma(-m+n)\Gamma(\frac{-m^2+z}{z+m})}{n\Gamma(-m+1)\Gamma(\frac{-m^2+nm+nz-m}{z+m})} + \frac{-m+n-1}{n}$$

Notando que $z = \frac{\hat{z}m}{m+1-\hat{z}}$, se puede mostrar que $P_n(\hat{z}(z))$ viene dado por:

$$P_n(\hat{z}(z)) = \frac{-m+n-1}{n} - \frac{(-m+\hat{z}-2)!(-m+n-1)!}{n(m-2)!(-m+n+\hat{z}-2)!}$$

Por el otro lado, se definir $\hat{P_n}(z)$ tal que:

$$\hat{P_n}(z) = \frac{\bar{B_n}(z)-\bar{B_n}(0)}{1-\bar{B_n}(0)}$$

En particular, dado que $\bar{B}_n(z) = \frac{\binom{m}{n}}{\binom{m+1-\bar{z}}{n}}$ y $\bar{z} = z\frac{m+1}{m+z}$, es decir, $z = \frac{m\hat{z}}{m+1-\hat{z}}$, se tiene que:

$$\hat{P}_n(\hat{z}(z)) = \frac{-m+n-1}{n} + \frac{(m-n-\hat{z}-1)!(m+1)!}{n(m-n)!(m+1-\hat{z})!}$$

Finalmente, se puede definir $D_n(\hat{z}(z))$ tal que:

$$D_n(\hat{z}(z)) = \hat{P}_n(\hat{z}(z)) - P_n(\hat{z}(z))$$

Desarrollando la expresión se tiene que:

$$D_n(\hat{z}(z)) = \frac{(m-n-\hat{z}-1)!(m+1)!}{n(m-n)!(m+1-\hat{z})!} + \frac{(-m+\hat{z}-2)!(-m+n-1)!}{n(m-2)!(-m+n+\hat{z}-2)!}$$

Se puede mostrar que si se cumple que $n \in \mathbb{N}$, $m > 0$ y $m \geq n$, se tiene que:

$$D_n(\hat{z}(z)) = 0$$

Es decir, se cumple que:

$$P_n(z) = \frac{\bar{B}_n(z) - \bar{B}_n(0)}{1 - \bar{B}_n(0)}$$

ANEXOS

(1) Se adjunta archivo con los cálculos realizados en Maple.