Probabilidad: Relación entre Distribución Hipergeométrica y de Bernoulli *

García Prado, Sergio
sergio@garciparedes.me

20 de septiembre de 2017

1. Distribución Hipergeométrica

La distribución hipergeométrica $X \sim H(n, N, D)$ se corresponde con una distribución de probabilidad para variables aleatorias discretas. Esta modeliza la situación de obtener k casos favorables, de n observaciones, sobre una población de N individuos, donde D representan la situación de éxito y $N - D$ la de fracaso. Por tanto, modeliza la misma situación que la distribución binomial, solo que en este caso se presupone que cada observación se realiza **sin reemplazamiento**, mientras que en la distribución binomial se asume reemplazamiento.

En las siguientes secciones se demuestra que una distribución hipergeométrica $X \sim H(n, N, D)$ es equivalente a la suma de n distribuciones de bernoulli dependientes entre sí. Por tanto, $X = X_1 + ... + X_i + ... + X_n$ donde $X_i \sim B(\frac{D}{N})$ representa la i-ésima observación.

2. Demostración para $n = 2$

A continuación se demuestra el caso para dos variables:

$$
\begin{aligned}
& X \sim H(2, N, D) && && (1) \\
& X = X_1 + X_2 && && (2) \\
& X_1 \sim B(\frac{D}{N}) && P[X_1 = 1] = \frac{D}{N} && (3) \\
& X_2 \sim B(\frac{D}{N}) && P[X_2 = 1] = P[X_2 = 1 \wedge X_1 = 0] + P[X_2 = 1 \wedge X_1 = 1] && (4) \\
& && = P[X_2 = 1 \mid X_1 = 0] * P[X_1 = 0] + P[X_2 = 1 \mid X_1 = 1] * P[X_1 = 1] && (5) \\
& && = \frac{D}{N-1} * (1 - \frac{D}{N}) + \frac{D-1}{N-1} * \frac{D}{N} && (6) \\
& && = \frac{D}{N} && (7)
\end{aligned}
$$

3. Demostración para $n = 3$

A continuación se demuestra el caso para dos variables:

*URL: https://github.com/garciparedes/probability-hypergeometric-to-bernoulli

$$
\begin{align}
& X \sim H(3, N, D) \tag{8} \\
& X = X_1 + X_2 + X_3 \tag{9} \\
X_1 \sim B(\frac{D}{N}) \qquad & P[X_1 = 1] = \frac{D}{N} \tag{10} \\
X_2 \sim B(\frac{D}{N}) \qquad & P[X_2 = 1] = P[X_2 = 1 \wedge X_1 = 0] + P[X_2 = 1 \wedge X_1 = 1] \tag{11} \\
& = P[X_2 = 1 \mid X_1 = 0] * P[X_1 = 0] + P[X_2 = 1 \mid X_1 = 1] * P[X_1 = 1] \tag{12} \\
& = \frac{D}{N-1} * (1 - \frac{D}{N}) + \frac{D-1}{N-1} * \frac{D}{N} \tag{13} \\
& = \frac{D}{N} \tag{14} \\
X_3 \sim B(\frac{D}{N}) \qquad & P[X_3 = 1] = ... \tag{15} \\
& = ... \tag{16} \\
& = \frac{D}{N} \tag{17} \\
& \tag{18}
\end{align}
$$

4. Demostración caso general

[TODO]

Referencias

[TG18] Jesús Alberto Tapia García. Muestreo Estadístico 1, 2017/18.