Distribuição Hipergeométrica

Essa distribuição é adequada quando consideramos extrações casuais feitas sem reposição de uma população dividida segundo dois atributos.

Para ilustrar, considere uma população de N objetos, r dos quais têm o atributo A e N - r têm o atributo B. Um grupo de n elementos é escolhido ao acaso, sem reposição. Estamos interessados em calcular a probabilidade de que esse grupo contenha k elementos com o atributo A. Pode-se ver facilmente, utilizando o princípio multiplicativo, que essa probabilidade é dada por

$$p_k = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}},$$

em que $\max(0, n - N + r) \leq k \leq \min(r, n)$.

Os pares (k, p_k) constituem a distribuição hipergeométrica de probabilidades.

notação: $\boldsymbol{X} \sim \mathbf{hip}(\boldsymbol{N}, \boldsymbol{r}, \boldsymbol{n})$.

Se definirmos a v.a. X como sendo o número de elementos na amostra que têm o atributo A, então $P(X = k) = p_k$.

Exemplo Em problemas de controle de qualidade, suponha que num lote de N = 100 peças, r = 10 sejam defeituosas. Escolhendo n = 5 peças sem reposição,

a) a probabilidade de não se obter peças defeituosas é

$$p_0 = \frac{\binom{10}{0}\binom{90}{5}}{\binom{100}{5}} = \frac{\binom{90}{5}}{\binom{100}{5}} \simeq 0,584,$$

b) a probabilidade de se obter pelo menos uma defeituosa é

$$p_1 + p_2 + \ldots + p_5 = 1 - p_0 \simeq 0,426.$$

$$E(X) = np,$$

$$\mathrm{Var}(X) = np(1 - p)\frac{N - n}{N - 1},$$

$$p = r/N$$

Observação: Se N for grande, quando comparado com n, então extrações com ou sem reposição serão praticamente equivalentes,

$$\boldsymbol{p_k \simeq b(k; n, p).}$$