7月23日「データ解析のための数理統計入門」第4章 命題4.21

2024-07-23

目次

命題 **0.1** (命題 4.21): 確率変数$X_1, ..., X_m$が独立で、各X_iがポアソン分布$Po(\lambda_i)$に従うとする。$X_1 + \cdots + X_m = n$という条件のもとで、$X_1, ..., X_m$ の条件付き分布は多項分布$Mult_m(n, p_1, ..., p_m)$ に従う。ただし、$p_i = \lambda_i / (\lambda_1 + \cdots + \lambda_m)$である。

証明.

$$Y = X_1 + \cdots + X_m, \quad \theta = \lambda_1 + \cdots + \lambda_m$$

とおくと、ポアソン分布の再生性により$Y \sim Po(\theta)$ となる。

$$P(X_1 = x_1, ..., X_m = x_m) = \prod_{i=1}^{m} \frac{\lambda_i^{x_i} e^{-\lambda_i}}{x_i!}$$

$$P(Y = n) = \frac{\theta^n}{n!} e^{-\theta}$$

条件付き確率は以下のようになる。

$$P(X_1 = x_1, ..., X_m = x_m \mid Y = n) = P(X_1 = x_1, ..., X_m = x_m, Y = n) / P(Y = n)$$

$P(X_1 = x_1, ..., X_m = x_m, Y = n)$ の同時確率は、$P(X_1 = x_1, ..., X_m = x_m)$が$x_1 + \cdots + x_m = n$ となるときである。

よって条件付き分布は多項分布$Mult_m(n, p_1, ..., p_m)$に従う。