General Bernoulli candidate scheme for masked data

Alex Towell

1/22/2022

We consider a general structure for candidate sets. Let $X_1, \ldots, X_m$ denote the Boolean random variables and let C be the random candidate set defined as

$$C = \{j \in \{1, \ldots, m\} : (X_j = 1)\}.$$

We consider the family of conditional independent Bernoulli distributions of X_j given $T_1 = t_1, \ldots, T_m = t_m$ for $j = 1, \ldots, m$. The conditional distribution of $X_1, \ldots, X_m$ given $T_1 = t_1, \ldots, T_m = t_m$ has a joint pdf given by

$$f(x_1, \ldots, x_m; \theta | t_1, \ldots, t_m) = \prod_{j=1}^{m} (p_j(t_1, \ldots, t_m, \theta))^{x_j} (1 - p_j(t_1, \ldots, t_m, \theta))^{1-x_j}. \tag{1}$$

Proof. By the Bernoulli assumption, the conditional pdf of X_j given $T_1 = t_1, \ldots, T_m = t_m$ may be written as

$$f(x_j; \theta | t_1, \ldots, t_m) = (p_j(t_1, \ldots, t_m, \theta))^{x_j} (1 - p_j(t_1, \ldots, t_m, \theta))^{1-x_j}.$$

By the assumption of mutual independence, the conditional joint pdf of $X_1, \ldots, X_m$ given $T_1 = t_1, \ldots, T_m = t_m$ is the product

$$f(x_1, \ldots, x_m; \theta | t_1, \ldots, t_m) = \prod_{i=1}^{m} f(x_j; \theta | t_1, \ldots, t_m).$$

Substituting the definition of $f(x_j; p_j, \theta | t_1, \ldots, t_m)$ into the above equation obtains the result

$$f(x_1, \ldots, x_m; \theta | t_1, \ldots, t_m) = \prod_{j=1}^{m} (p_j(t_1, \ldots, t_m, \theta))^{x_j} (1 - p_j(t_1, \ldots, t_m, \theta))^{1-x_j}.$$

$\square$

One of the most common Bernoulli models is given by making X_j be a function of the component with the minimum time-to-failure. In this case,

$$f(x_k | t_1, \ldots, t_m) = \begin{cases} 1 & \text{if } x_k = 1 \text{ and } t_k < t_j \text{ for all } j \neq k \\ 0 & \text{if } x_k = 0 \text{ and } t_k < t_j \text{ for } j \neq k \\ \gamma_k & \text{if } x_k = 1 \text{ and } t_k > t_j \text{ for any } j \neq k \\ 1 - \gamma_k & \text{if } x_k = 0 \text{ and } t_k > t_j \text{ for any } j \neq k. \end{cases}$$

Observe that in the main paper, we define the random variable K as the component label with the minimum time-to-failure. Thus, we can rewrite the above as

$$f(x_j | k) = \begin{cases} 1 & \text{if } x_j = 1 \text{ and } j = k \\ 0 & \text{if } x_j = 0 \text{ and } j = k \\ \gamma_k & \text{if } x_j = 1 \text{ and } j \neq k \\ 1 - \gamma_k & \text{if } x_j = 0 \text{ and } j \neq k. \end{cases}$$