Chapter 7

Perfect aggregation in reliability analysis

7.1 Introduction

In this chapter, we are interested in finding joint prior distributions for the system state probabilities in reliability systems so that the posterior system failure probability obtained by updating this prior with component-level data is the same as if we instead used system-level data. This property has been described as "perfect aggregation" (Azaiez, 1993).

More specifically, a *Bernoulli system* is a coherent system made up of components, each of which can either succeed or fail according to a Bernoulli process. See Azaiez (1993) for a more rigorous definition. One should note that Azaiez generally restricted himself to systems with no replications (i.e., systems that "can be represented, using only and/or logic, in such a way that each component appears only once, which for example excludes k-out-of-n systems for k > 1"). However, Azaiez relaxed this restriction at the end of his thesis, where he considered systems with dependent failures that may in particular have replicated components. Here, we consider only Bernoulli systems with dependent failures; i.e., we start where Azaiez left off.

7.2 The problem

Consider a Bernoulli system made up of N components. Adopting the notation of Azaiez (1993), let

$$\Omega = \{\omega = (\omega_1, \ldots, \omega_N) \mid \omega_i \in \{0,1\} \quad \forall\, i = 1,2,\ldots,N\}: \tag{7.1}$$

be the set of all possible N-dimensional binary vectors,

E_{ω}: the event that, for all *i*, component *i* fails if $\omega_i = 0$ and succeeds if $\omega_i = 1$, for all ω in Ω,

p_{ω}: the probability of event E_{ω}, for all ω in Ω,

From the above we have:

$$\sum_{\omega \in \Omega} p_{\omega} = 1 \qquad (7.2)$$

So, the set $\{p_{\omega} | \omega \in \Omega\}$ is a composition. Also, note that the system failure probability is given by

$$P_f = \sum_{\omega \in \Psi} p_{\omega}\,, \qquad (7.3)$$

where $\Psi \subset \Omega$ is the set of states that cause the system to fail.

Given a joint prior distribution for the system state probabilities, $f(p_{\omega} : \omega \in \Omega)$, let the term "disaggregate analysis" refer to the process of first updating this prior distribution using data on the occurrences of the individual system states, and then aggregating the resulting posterior to obtain a distribution for the system failures probability P_f. Similarly, let "aggregate analysis" refer to the process of first aggregating the prior distribution $f(p_{\omega})$ to obtain a prior distribution for the system failure probability P_f, and then updating this distribution using only data about the successes and failures of the system as a whole. If the distributions for the system failures probability resulting from the aggregate and disaggregate analyses agree, then we say "perfect aggregation" has been attained.

As stated earlier, we are interested in finding joint prior distributions for the system state probabilities p_{ω} so that perfect aggregation holds. The purpose is to help analysts determine when they can safely disregard the disaggregate data on the occurrences of the individual system states.

Assume that we have data from K_0 tests of the system, during which event E_ω occurred exactly k_ω times for all ω in Ω, where $K_0 = \sum_{\omega \in \Omega} k_\omega$. Then the disaggregate data on the system state occurring in each trial is given by

$$DD = \{k_\omega | \omega \in \Omega\} \tag{7.4}$$

and the number of system failures is given by $k^* = \sum_{\omega \in \Psi} k_\omega$. Similarly, *AD*={ K_0, k^* } is the aggregate data on the successes and failures of the system as a whole.

Let the number of elements in Ω be n+1, and let the number of elements in Ψ be m. Relabel the elements in Ψ by 1,2,...,m, and the elements in $\Omega \setminus \Psi$ by m+1,m+2,...,n+1. Hence, p_i is the probability of the event E_i, $i = 1,2,\ldots,n+1$. Finally, let $P = (p_1, p_2, ..., p_n)$.

Note that perfect aggregation concept is equivalent to saying that the function $T(\omega_1,\ldots,\omega_{n+1}) = (K_0, k^*)$ is a minimal sufficient statistic for P_f. For this terminology see Dudewicz and Mishra (1988).

7.3 Hazen's conditions for perfect aggregation

In this section we briefly describe some unpublished work by Hazen (1992) on necessary and sufficient conditions for perfect aggregation. Related work can be found in Rubin (1976), illustrating that the problem of perfect aggregation can be interpreted as a problem of missing data.

Let *DD* and *AD* represent the disaggregate and aggregate data, respectively, and let the distributions of *DD* and *AD* depend on the vector of unknown system state probabilities *P*. Suppose there are some subvectors δ and γ of parameters such that: (1) *P* can be expressed as (δ, γ) and vice versa; (2) given δ, the *AD* is independent of γ; and (3) given (*AD*, γ), *DD* is independent of δ. In other words,

$$AD \mid (\delta,\ \gamma) \overset{d}{=} AD \mid \delta \tag{7.5}$$

$$DD \mid (AD,\ \delta\ ,\ \gamma) \overset{d}{=} DD \mid (AD,\ \gamma) \tag{7.6}$$

where $AD \mid \delta$ is the conditional probability distribution of the aggregate data given a particular value of δ, and so on.

We say that a set of functions A includes all the "powers" of γ if the function defined by

$$\gamma \xrightarrow{H} \gamma_1^{k_1} \ldots \gamma_{m-1}^{k_{m-1}} \left(1 - \sum_{i=1}^{m-1} \gamma_i\right)^{k^* - k_1 - \ldots - k_{m-1}} \gamma_m^{k_{m+1}} \ldots\ \gamma_{n-1}^{k_n} \left(1 - \sum_{i=m}^{n-1} \gamma_i\right)^{k_{n+1}}$$

is a member of A for all possible k_i (i=1,...,n+1) such that $\sum_{i=1}^{n+1} k_i = K_0$ and all possible k^* such that $0 \le k^* \le K_0$.

Hazen states the following conclusion:

> *If conditions (7.5) and (7.6) hold, then for perfect aggregation to hold it is sufficient that* δ , γ *be independent. If in addition the set A of "link" functions*
>
> $$A = \{H: \gamma \xrightarrow{H} DD \mid (AD,\ \gamma)\}$$

includes all the "powers" of γ*, then* δ, γ *being independent is necessary for perfect aggregation to hold.*