Chapter 4.
MEASUREMENTS SELECTION CRITERION

Sometimes, more than one auxiliary measurements are taken into consideration, and there are four in our study. Different criteria can be applied to choose the best one among all candidates depending on different purposes. For example, if the goal is to make inferences about the lifetime distribution, minimizing the variance of $\widehat{\alpha_0}$ (the baseline hazard) is suggested. We do not recommend those criteria calculated for the information matrix which contains the parameters appearing in degradation paths such as the trace or determinant, because it is affected by different scales of Y. In this chapter, we propose a new criterion, CCP (correct classification probability) to select the best auxiliary measurement among all the candidates.

4.1 General Concepts of the CCP

A good degradation measurement should have a strong relationship with the lifetime. At time t, if the goodness of a tested unit is judged by whether its corresponding observation of degradation measurement, Y_t, is above or below the threshold, τ, then the conditional probability of correct classification given the lifetime T, which is defined as

$$CCP_{\boldsymbol{\theta}}(t|T) = \begin{cases} P_{\boldsymbol{\theta}}(Y_t < \tau|T) & \text{if} \quad t \geq T \\ P_{\boldsymbol{\theta}}(Y_t > \tau|T) & \text{if} \quad t < T, \end{cases}$$

should be high for all T, ideally. Considering the average and taking the expectation over T with respect to the lifetime distribution $F(T;\boldsymbol{\alpha})$, CCP is defined as:

$$CCP_{\boldsymbol{\theta}}(t) = \mathrm{E}_{\boldsymbol{\theta}}[CCP_{\boldsymbol{\theta}}(t|T)].$$

It is the probability of an inspected item being classified as an active or a failed one according to if the observation of the measurement Y_t is above or below the threshold. Recall that $\boldsymbol{\theta} = (\alpha, \beta_0, \beta_1, \sigma^2)$. When CCP is chosen to be the selection criterion, we look for an auxiliary measurement with the highest CCP.

The CCP is a general device. It can be used as a selection criterion among any kind of degradation measurements with any underlining lifetime distribution at any inspecting time by comparing the CCP directly. However, it can be rarely computed exactly. The most straightforward way of applying the CCP is to estimate it by using the Monte Carlo methods. Assume that at a inspecting time t, an observation for a degradation measurement is Y_t. Further, the lifetime T is with the density function $f_{\boldsymbol{\alpha}}(T)$, and the conditional density of Y_t given T is $f_{\boldsymbol{\theta}}(Y|T)$. Then, clearly

$$CCP_{\boldsymbol{\theta}}(t) = P_{\boldsymbol{\theta}}(Y_t < \tau, T < t) + P_{\boldsymbol{\theta}}(Y_t > \tau, T > t).$$

and the corresponding Monte Carlo approximation can naturally be done by the procedure below. First, draw B random samples $(T_1, Y_1), \cdots, (T_B, Y_B)$ according to

$$f_{\boldsymbol{\theta}}(Y, T) = f_{\boldsymbol{\alpha}}(T) f_{\boldsymbol{\theta}}(Y|T).$$

Then, the CCP can be approximated by

$$CCP_{\boldsymbol{\theta}}(t) \approx \frac{1}{B}\sum_{b=1}^{B}[I(Y_b < \tau, T_b < t) + I(Y_b > \tau, T_b > t)].$$

In practice, the parameters are unknown. The CCP can be estimated by replacing the true parameters with their estimators in the above Monte Carlo method.

4.2 The CCP to the Linear Degradation Model

Here we are going to illustrate some nice properties of the CCP, under the special model assumptions. Recall that if T were known the degradation measure at time t is $\beta_0+\beta_1\frac{t}{T}+\epsilon$. The observation of the measurement, Y_t, would follow a normal distribution with mean $\beta_0+\beta_1\frac{t}{T}$ and variance σ^2. Thus, the probability of Y_t being above or below the threshold would depend on both the position of the mean degradation process, $\beta_0+\beta_1\frac{t}{T}$, and the random error, σ^2.

Definition: When hazard rate of the lifetime distribution is α_0 and the inspecting time is t, the best degradation measurement, $Y^*(t,\alpha_0)$, is the one with parameter $\boldsymbol{\theta}^*$ which satisfies

$$CCP_{\boldsymbol{\theta}^*}(t) = \max_{Y\in\mathcal{D}} CCP_{\boldsymbol{\theta}}(t),$$

where $\mathcal{D}$ is the set of candidates of degradation measurement.

Note that the change of Y will change the parameter $\boldsymbol{\theta}$, i.e. different Y is with different $(\beta_0,\beta_1,\sigma^2)$ but α_0 is fixed. Hence, $\boldsymbol{\theta}=\boldsymbol{\theta}(Y)$.

The following theorem shows that the choice of the best degradation measurement is irrelevant to both t and α_0.

Theorem 1 *Fix t, $CCP_{\boldsymbol{\theta}}(t)$ increases from 0.5 to 1 as $|\beta_1|/\sigma$ increases from 0 to ∞. That is the best degradation measurement is the one with the largest $|\beta_1|/\sigma$.*

Proof. According to model assumptions, we have $Y_t \sim N(\beta_0+\beta_1\frac{t}{T}, \sigma^2)$ and $\tau = \beta_0+\beta_1$, and this implies

$$CCP_{\boldsymbol{\theta}}(t|T) = \begin{cases} \Phi(\frac{\beta_1}{\sigma}(1-\dfrac{t}{T})) & \text{if} \quad t \geq T, \\ \Phi(\frac{\beta_1}{\sigma}(\dfrac{t}{T}-1)) & \text{if} \quad t < T, \end{cases} \tag{4.1}$$

where $\Phi(.)$ is the standard normal distribution function. Therefore

$$\begin{aligned} CCP_{\boldsymbol{\theta}}(t) &= \mathrm{E}_{\boldsymbol{\theta}}[CCP_{\boldsymbol{\theta}}(t|T)] \\ &= \int_t^{\infty} f(T;\alpha_0)\Phi(\frac{\beta_1}{\sigma}(\frac{t}{T}-1))dT + \int_0^t f(T;\alpha_0)\Phi(\frac{\beta_1}{\sigma}(1-\frac{t}{T}))dT. \end{aligned}$$

When $t \geq T$, $\frac{\beta_1}{\sigma}(\frac{t}{T}-1) > 0$ and when $t \geq T$, $\frac{\beta_1}{\sigma}(1-\frac{t}{T}) > 0$ $(\beta_1 < 0)$. Hence both of the integrands increase in $|\beta_1|/\sigma$. Because

$$\int f(T;\alpha_0)\Phi(.)dT < \int f(T;\alpha_0)dT < \infty,$$

the upper and lower bounds can be obtain by bounded converge theorem and letting $|\beta_1|/\sigma$ approaching 0 and ∞, respectively.

When $\beta_1 > 0$, changing the signs of the inequalities in the definition of $CCP_{\boldsymbol{\theta}}(t|T)$, the result can be derived similarly and the theorem is proved.

Theorem 1 states that the best degradation measurement is uniform in both t and α_0. That means $Y^*(t,\alpha_0)$ is independent on both of its coordinates. Therefore we can eliminate t and α_0, and denote it as Y^* for simplicity.

Remark 1 Since the variance σ^2, is the characteristic of misclassification, it can be regarded as a parameter of "noise". Under the same variance, the trend can be obtained more easily when the slope β_1, is large in its absolute value. Thus, $|\beta_1|$ can be regarded as a parameter of "signal". The smaller σ^2 and the larger $|\beta_1|$ cause the larger CCP. Under the same lifetime distribution, the CCP's are affected by the ratio of "signal" and "noise", $r = |\beta_1|/\sigma$.

Not surprisingly, the result coincides with the choice of the best SN ratio in signal-response systems (cf. Chapter 11 of Wu and Hamada, 2000), or dynamic SN ratio in Taguchi (1991). The choice of the best degradation can be regarded as the choice of the best measurement system which is a kind of the signal-response systems. The lifetime can be considered as the unobserved input of the system, the observation Y is the response and the performance measurement is $|\beta_1|/\sigma$. If the criterion is to find the system which is with the highest correct probability to classify tested units into two categories according to the threshold, then the best system is the one with the highest $|\beta_1|/\sigma$. In Taguchi (1991), the goal is how to estimate the unknown input with the least variance, while in our case the $\frac{1}{T}$. Solving the linear system

$$Y_t = \beta_0 + \beta_1 \frac{t}{T},$$

the estimator of $\frac{1}{T}$ is given by

$$\widehat{\frac{1}{T}} = \frac{Y_t - \beta_0}{\beta_1 t},$$

and the variance is

$$Var[\widehat{\frac{1}{T}}] = \frac{\sigma^2}{\beta_1^2 t^2}.$$

Therefore, the best selection is the one with the highest $|\beta_1|/\sigma$. Both criteria involve the accuracy of T's estimator. The more accurately of T estimated, the less variance and the higher CCP obtained.

Next theorem characterizes a curvature property of the $CCP_{\boldsymbol{\theta}}(t)$.

Theorem 2 *There exists a* $t^* = t^*(\boldsymbol{\theta})$ *such that* $CCP_{\boldsymbol{\theta}}(t)$ *is decreasing in* $(0, t^*]$ *and increasing in* (t^*, ∞).

Proof. For a given $\boldsymbol{\theta}$, by using Leibitz's rule, we obtain

$$\frac{\partial CCP_{\boldsymbol{\theta}}(t)}{\partial t} = \int_t^{\infty} f(T;\alpha_0)\frac{\beta_1}{\sigma T}\phi(\frac{\beta_1}{\sigma}(\frac{t}{T}-1))dT - \int_0^t f(T;\alpha_0)\frac{\beta_1}{\sigma T}\phi(\frac{\beta_1}{\sigma}(1-\frac{t}{T}))dT.$$

It is easy to verify that both $\frac{\beta_1}{\sigma T}\phi(\frac{\beta_1}{\sigma}(\frac{t}{T}-1))$ and $\frac{\beta_1}{\sigma T}\phi(\frac{\beta_1}{\sigma}(1-\frac{t}{T}))$ in their integration domain are finite, so the integrations exist. In addition, both of the integrands are positive. Hence, there exists a t^* such that $\frac{\partial CCP(t,\boldsymbol{\theta})}{\partial t} \leq 0$ as $t \leq t^*$ and > 0 as $t > t^*$, which completes the proof.

The CCP's with different combinations of parameters are displayed in Figure 1.

Remark 2 The effects of lifetime distributions to the CCP can be explained from $CCP_{\boldsymbol{\theta}}(t|T)$. It is easy to see from (4.1) that the $CCP_{\boldsymbol{\theta}}(t|T)$ decreases until reaches the minimum 0.5 at $t = T$, and then increases, which means if the lifetime was given, the maximum of misclassification probability happens when the mean degradation path reaches the threshold. The $CCP_{\boldsymbol{\theta}}(t)$ is the average of $CCP_{\boldsymbol{\theta}}(t|T)$, so it inherits the pattern from $CCP_{\boldsymbol{\theta}}(t|T)$. When the mean of a lifetime distribution is large, the degradation path will reach the threshold latter, in the average sense. Hence, the minimum of CCP happens earlier as α_0 is larger.

In practice, $\boldsymbol{\theta}$ is not available. Therefore, we choose the measurement with the

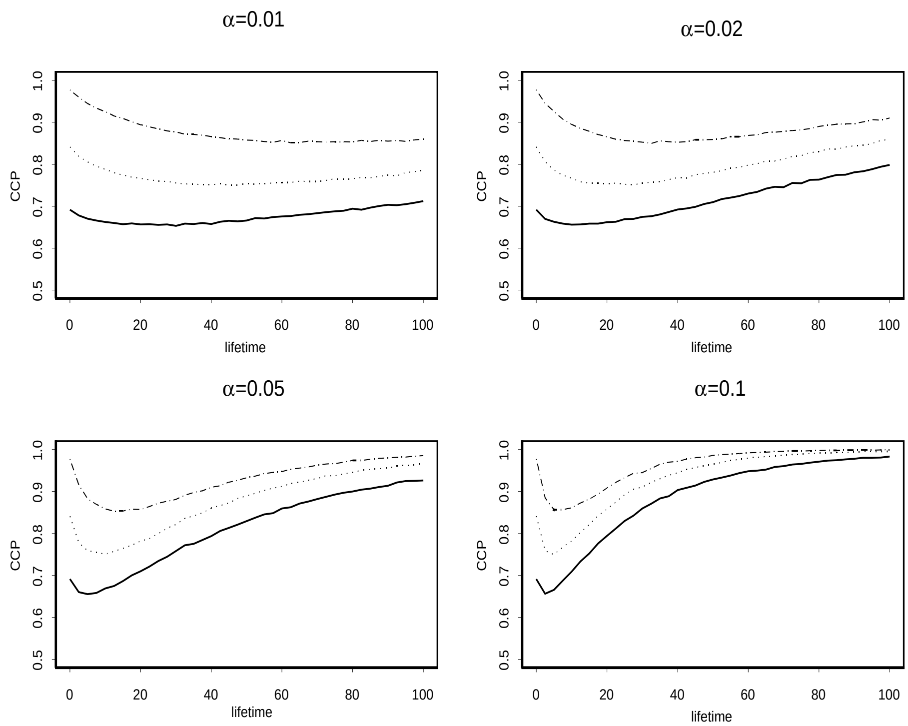

Figure 1: CCP's in different parameter combinations (—- for $r = 0.5$, $\cdots\cdot$ for $r = 1.0$ and $\cdot$—$\cdot$ for $r = 2$).

highest $|\widehat{\beta_1}|/\widehat{\sigma}$ as the best degradation measurement. The accuracy can be evaluated by its standard deviation. This will be explained more details in Chapter 5.