Binomial and Sequential Testing Reference: pp. 342-348 of Ebeling(2nd)

Maghsoodloo

Ebeling provides an example of a single sampling plan, pp. 342-343, in RE testing where n units are tested for their TsTF and the number of failed units, N_f, by end of test time, T, is recorded as r. The objective is to determine the two decision variables n and N_f = r such that the producer's risk, α = Pr of type I error at $R = R_0$, is roughly 0.05, and the consumer's risk, β = Pr of type II error, say is roughly 8% when $R = R_1 < R_0$. Note that in his Table 13.2, p. 343, his R_1 = my R_0 and his R_2 = my R_1. I prefer 0 and 1 because of H_0 and H_1. Since N_f has a binomial distribution, under H_0: $R = R_0$, the Pr of accepting a submitted lot of size n for static RE testing is given by

P_a(Good lot)= B(r; n, R_0) = $\sum_{i=0}^{r} {}_nC_i R_0^{n-i} Q_0^i \geq 1-\alpha$ = Pr of accepting a good lot, P_a(Bad lot)= B(r; n, R_1) = $\sum_{i=0}^{r} {}_nC_i R_1^{n-i} Q_1^i \leq \beta$ = The Pr of accepting a bad lot. As an example, suppose an $R_0 = 0.98$ is desired with $\alpha = 0.05$ but when lots are as bad as $R_1 = 0.92$, we wish to reject them with a Pr of roughly 0.91 (or Consumer's risk = 0.09). Then, we need to solve the two equations

$\sum_{i=0}^{r} {}_nC_i 0.02^i 0.98^{r-i} \geq 0.95$ and $\sum_{i=0}^{r} {}_nC_i (0.08)^i (0.92)^{r-i} \leq 0.09$. I have provided an Excel file on my website that solves for possible values of n and r. My Excel file shows that n = 130 and r = 6 come close to meeting the above requirements. If an Excel file is not available, then

$\sum_{i=0}^{r} {}_nC_i 0.02^i 0.98^{r-i} \cong \sum_{i=0}^{r} \frac{(0.02n)^i}{i!} e^{-0.02n} \geq 0.95$, and $\sum_{i=0}^{r} {}_nC_i R_1^{n-i} Q_1^i \cong \sum_{i=0}^{r} \frac{(0.08n)^i}{i!} e^{-0.08n} \leq \beta$ can be solved to obtain approximate solutions for n and r. For the given plan (130, r= 6), this Poisson approximation to the Binomial gives $1-\alpha$ =0.9828, while β =0.1069.

We will discuss, as does Ebeling on his p. 345, the one-sided procedure for a sequential sampling plan, which is the most common in RE testing of the type $H_0: \lambda = \lambda_0$ versus the alternative $H_1: \lambda > \lambda_0$. In the one-sided procedure, the alternative value of process failure rate is denoted by $\lambda_1 > \lambda_0$ (a right-tailed test). In carrying out sequential testing, generally samples of size n (= 1, but not always) is taken at each stage of the sampling plan, the sequential test statistic (SQTS) is updated, and a 3-way decision is made as follows: (1) accept H_0, (2) reject H_0, or (3) continue the experiment by taking additional observation(s). I would venture to say, that it is rare to take differing sample sizes at different stages in RE Engineering, although this is sometimes done in acceptance sampling for FNC =fraction nonconforming (especially in a 2-stage sampling plan where $n_2 = 2\times n_1$ but then in this case the decision about a submitted lot is definitely made after stage 2 because the difference between the acceptance and rejection numbers at the 2^{nd} stage is always equal to 1).

For example, the double sampling plan $n_1 = 50$, $c_1 = 1$, $n_2 = 100$, and $c_2 = 4$ implies that a random sample of size 50 is taken from a large lot (usually N > 1000 units) and the lot is immediately accepted if the number of defectives, D_1, at stage 1 is ≤ 1 and the lot is immediately rejected at stage 1 if $D_1 > 4$. If the lot quality is mediocre at stage 1, where $D_1 = 2, 3$, or 4, then the lot is given a second chance by taking an additional sample of size $n_2 = 100$. At the 2^{nd} stage, the lot is accepted only if the total number of defectives in the 150 sampled units is 4 or less, and it is rejected at stage 2 only if $D_1 + D_2 > 4$.

When sequential sampling is applied in SPC or RE Engineering, first the reader should be cognizant of the fact that the inspection is not always by attributes (like on p. 347 of Ebeling) and that the decision about the MTTF or failure of a Weibull (or Exponential) process is carried out with a statistic (like TTF) whose range space is on a continuous (or dense) scale. The very basic statistical idea behind testing the hypothesis H_0: $\lambda = \lambda_0$ versus the alternative $H_1 : \lambda > \lambda_0$ is to make use of the Neyman-Pearson Lemma by using the likelihood ratio statistic

$$y_r = \frac{\prod_{i=1}^{r} f(t_i;\lambda_1)dx_i}{\prod_{i=1}^{r} f(t_i;\lambda_0)dx_i} = \frac{\prod_{i=1}^{r} f(t_i;\lambda_1)}{\prod_{i=1}^{r} f(t_i;\lambda_0)} = \frac{P_a(\lambda = \lambda_1)}{P_a(\lambda = \lambda_0)} \qquad (121)$$

In equation (121) the statistic in the numerator on the RHS is related to the occurrence Pr of the sample ($t_1, t_2, ..., t_r$) given that $\lambda = \lambda_1 > \lambda_0$ because the likelihood (or Pr) of obtaining such a sample is given by $[f(t_1;\lambda)dt_1]\times[f(t_2; \lambda_1)dt_2]\times ... \times [f(t_r; \lambda_1)dt_r] = \prod_{i=1}^{r}\left[f(t_i;\lambda_1)\,dt_i\right]$, where $f(t;\lambda)$ represents the failure density of the sampled population. The quantity $\prod_{i=1}^{r} f(t_i;\lambda_1)$ in this last Pr statement is called the likelihood function under H_1, and after the sample is drawn and the sample values $t_1, t_2, ..., t_r$ are known numbers (no longer rvs), then the likelihood function $\prod_{i=1}^{r} f(t_i;\lambda_1)$ is only a function of the parameter λ, which is denoted by $L(\lambda) = \prod_{i=1}^{r} f(t_i;\lambda)$. In other words, the numerator of (121) gives the likelihood that λ lies in an extremely small interval around λ_1 and its denominator provides the Pr that λ lies in an infinitesimally small interval around λ_0. Therefore, when y_r is very large, we will reject H_0 in favor of H_1 at the stage r (or at the time t_r); when y_r is too small, we will favor H_0 over H_1, and when it is neither, then the sample point is still in the indecision region of sample space so that we have to continue on sampling to the stage r+ 1. Further, for notational convenience, let $\mathcal{L}(\lambda) = \ln[L(\lambda)]$, i.e., $\mathcal{L}(\lambda)$ is the natural logarithm of the likelihood function, as before. Working with the natural logarithm is easier than using likelihood function $L(\lambda)$ itself. Before providing such an example, we must provide some partial answer as to how large y_r should be before we adopt $\lambda = \lambda_1$ and how small y_r should be before we adopt $\lambda = \lambda_0$. In other words, we will reject H_0 if $y_r \geq B$, we will accept H_0 if $y_r \leq A$, and will continue sampling if the value of y_r lies in the indecision interval $A < y_r < B$, which depends on the specified values of the LOS α and the type II error Pr β. Note that in general $B >> A$.

One advantage that sequential testing has over the fixed-size sample tests is the fact that we can decide in advance on the sizes of type I and II errors (generally $\alpha \leq 0.05$ and $\beta \leq 0.10$) that we are willing to tolerate, which is unlike the fixed-size sampling where α is specified a priori and β is computed for different parameter values. Suppose we fix α at 0.05 and β at 0.10 for detecting a upward shift in a process failure rate; then clearly the value of B must be proportional to $1 - \beta$ (= the Pr of rejecting a false H_0). It was shown by A. Wald and G. A. Barnard (1945 and 1946) that the acceptance and rejection limits must satisfy $\frac{\beta}{1-\alpha} \leq A$ and $B \leq \frac{1-\beta}{\alpha}$, respectively. The statistic y_r is referred to as the sequential Pr ratio test (SPRT) with approximate boundaries $A \cong \frac{\beta}{1-\alpha}$ and $B \cong \frac{1-\beta}{\alpha}$. Note that the boundaries for the indecision interval is designed in such a manner that the true type I and II Prs of the sequential test are at most α and β, respectively. The reader should observe that the denominator of the SPRT statistic y_r = $\frac{\prod_{i=1}^{r} f(t_i;\lambda_1)\,dt_i}{\prod_{i=1}^{r} f(t_i;\lambda_0)\,dt_i}$ gives the likelihood Pr that H_0 is true, and therefore, this Pr (i.e., the denominator) must be directly proportional to 1 – α (= the Pr of accepting a true H_0) and inversely proportional to β (= the Pr of accepting a false H_0). This is how the expression for lower boundary A for the indecision interval $A < y_r < B$ was constituted.

Now consider the Example 13.8 on pp. 345-346 of Ebeling, where the objective is to develop a SPRT for a CFR, where $\lambda_0 = 0.00125$, $\alpha = 0.05$, and $\lambda_1 = 0.0014286$, $\beta = 0.10$. Because the underlying failure distribution is exponential, Eq. (121) becomes $y_r = \frac{\prod_{i=1}^{r} f(t_i;\lambda_1)\,dt_i}{\prod_{i=1}^{r} f(t_i;\lambda_0)\,dt_i} =$

$= \frac{\prod_{i=1}^{r} f(t_i;\lambda_1)}{\prod_{i=1}^{r} f(t_i;\lambda_0)} = \frac{\prod_{i=1}^{r} \lambda_1 e^{-\lambda_1 t_i}}{\prod_{i=1}^{r} \lambda_0 e^{-\lambda_0 t_i}}$, A =0.10/(0.95) =0.10526, and B = 0.90/.05 = 18. Thus, the indecision interval is given by $0.10526 \le \frac{\prod_{i=1}^{r} \lambda_1 e^{-\lambda_1 t_i}}{\prod_{i=1}^{r} \lambda_0 e^{-\lambda_0 t_i}} = \frac{\lambda_1^r e^{-\sum_{i=1}^{r} \lambda_1 t_i}}{\lambda_0^r e^{-\sum_{i=1}^{r} \lambda_0 t_i}} \le 18$. Taking ln of all terms we obtain, $-2.2513 \le r\ln(\lambda_1) - \sum_{i=1}^{r} \lambda_1 t_i - r\ln(\lambda_0) + \sum_{i=1}^{r} \lambda_0 t_i \le 2.8904$, or $-2.2513 \le$

$$r\ln(\lambda_1/\lambda_0) - (\lambda_1-\lambda_0)\sum_{i=1}^{r} t_i \le 2.8904 \rightarrow -2.2513\ \text{-}r\ln(\lambda_1/\lambda_0) \le -(\lambda_1-\lambda_0)\sum_{i=1}^{r} t_i \le 2.8904 -$$

$$r\ln(\lambda_1/\lambda_0) \rightarrow [-2.8904 + r\ln(\lambda_1/\lambda_0)]/(\lambda_1-\lambda_0)\,] \le \sum_{i=1}^{r} t_i = \tau_r \le [2.2513 + r\ln(\lambda_1/\lambda_0)]/$$

$$(\lambda_1-\lambda_0)\,] \rightarrow [-2.8904 + r\ln(\lambda_1/\lambda_0)]/(\lambda_1-\lambda_0) \le \tau_r \le [2.2513 + r\ln(\lambda_1/\lambda_0)]/(\lambda_1-\lambda_0)\,].$$

Substituting ln(0.0014286/0.00125)= 0.1336 and $\lambda_1 - \lambda_0$ = 0.00017857, we obtain τ_L = –2.8904 $+ r\ln(\lambda_1/\lambda_0)]/(\lambda_1-\lambda_0)$ = –16186.24 + 747.7758r, and τ_U = 12607.28+ 747.7758r. These are consistent with what Ebeling gives on his p. 346. Thus, the plan is to compute the value of $\tau_r = \sum_{i=1}^{r} t_i$ for r =1, 2 , 3, 4, … and test this against the limits τ_L and τ_U. If $\tau_r > \tau_U$, immediately accept the submitted lot; If τ_r is less than τ_L, immediately reject the lot; otherwise, test one more unit to failure.

Ebeling also applies the SQTS to binomial testing, where $H_0 : R = R_0$ versus the alternative $H_1: R < R_0$. Let y_n = equal to number of survivors out of n units on test. Then, for $y_{n,y}$ = y survivors the LRTS is given by $\frac{{}_nC_y R_1^y Q_1^{n-y}}{{}_nC_y R_0^y Q_0^{n-y}} = \frac{R_1^y Q_1^{n-y}}{R_0^y Q_0^{n-y}}$, where $R_1 < R_0$ is the alternative value of R under H_1, and as before A = β /(1–α) and B =(1–β)/α, and the most common values of α and β are 5% and 10%, respectively. As Ebeling, I will also go through his Example 13.9 on pp. 347-348. Here, R_0 = 0.90 and R_1 = 0.85, and α, β are the standard 5% and 10%, respectively. Then, A = 0.10/0.95 = 0.105263 and B = 18. Thus, we accept the lot iff $y_{n,r} = \frac{R_1^y Q_1^{n-y}}{R_0^y Q_0^{n-y}}$ < 0.105263 and reject the lot iff $y_{n,r}$ > 18. We will continue testing iff

$$0.105263 < y_{n,r} = \frac{R_1^y Q_1^{n-y}}{R_0^y Q_0^{n-y}} < 18 \qquad (122)$$

Or: $-2.2512918 < \ln(R_1^y Q_1^{n-y}) - \ln(R_0^y Q_0^{n-y}) < 2.890371757896165$

Or: $-2.2512918 < y\ln(R_1) + (n-y)\ln(Q_1) - y\ln(R_0) - (n-y)\ln(Q_0) < 2.890371758$

Or: $-2.2512918 < -0.462624y + n\ln(Q_1 / Q_0) < 2.890371758 \rightarrow$

Or: $-2.2512918 < -0.462624y + 0.405465n < 2.890371758 \rightarrow$

Or: $-2.2512918 - 0.405465n < -0.462624y < 2.890371758 - 0.405465n \rightarrow$

Or: $(-2.2512918 - 0.405465n)/(-0.462624) > \mathbf{y} > (2.890371758 - 0.405465n)/-0.462624 \rightarrow$

Or: $-6.247783826 + 0.87644724n < y < 4.8663583 + 0.87644724n$

Thus, the above plan dictates that we test one unit to failure at-a-time and if $N_f(n) < -6.247783826 + 0.87644724n$, we accept the lot; if $N_f(n) > 4.8663583 + 0.87644724n$, we immediately reject the lot. Otherwise, test another unit to failure.

As Ebeling states on his page 348, Eq. (122) can also be used for testing acceptable and reject-able cumulative repair-hazard functions $H(t^*) = P_0$ and $H(t^*) = P_1$, where R_1 in (122) is replaced with P_1 and R_0 is replaced with P_0. ASA, the LRTS $\frac{P_1^y Q_1^{n-y}}{P_0^y P_0^{n-y}}$ exceeds 18 reject the lot, and accept iff $\frac{P_1^y Q_1^{n-y}}{P_0^y P_0^{n-y}} < 0.105263$ ($\alpha = 0.05$ and $\beta = 0.10$), and continue obtaining repair data as long as $A < \frac{P_1^y Q_1^{n-y}}{P_0^y P_0^{n-y}} < B$.