Essentially ML ASN-Minimax double sampling plans

Eno Vangjeli

Abstract: Subject of this paper is ASN-Minimax (AM) double sampling plans by variables for a normally distributed quality characteristic with unknown standard deviation and two-sided specification limits. Based on the estimator p^* of the fraction defective p, which is essentially the Maximum-Likelihood (ML) estimator, AM-double sampling plans are calculated by using the random variables p_1^* and p_p^* relating to the first and pooled samples, respectively. Given p_1, p_2, α, and β, no other AM-double sampling plans based on the same estimator feature a lower maximum of the average sample number (ASN) while fulfilling the classical two-point condition on the corresponding operation characteristic (OC).

Keywords: Acceptance sampling by variables, ASN-Minimax double sampling plan, essentially Maximum-Likelihood estimator

1. Introduction

When carrying out sampling inspection for a normally distributed characteristic $X \sim N(\mu, \sigma)$, $\sigma > 0$ the following four cases arise:

(i) One-sided specification limit, σ known

(ii) Two-sided specification limits, σ known

(iii) One-sided specification limit, σ unknown

(iv) Two-sided specification limits, σ unknown.

In this paper, we deal with ASN-Minimax (AM) double sampling plans for case (iv). Let L be a lower and U an upper specification limit to X. The fraction defective function $p(\mu, \sigma)$ is defined as:

$$p(\mu, \sigma) := P(X < L) + P(X > U) = \Phi\left(\frac{L-\mu}{\sigma}\right) + \Phi\left(\frac{\mu-U}{\sigma}\right), \qquad (1)$$

where Φ denotes the standard normal distribution function. Note, $p(\mu, \sigma)$ is a three-dimensional function. For different levels of p, corresponding iso-p-lines arise symmetrically to $\mu_0 = \dfrac{L+U}{2}$ on the μ-σ-plane. A figure containing different iso-p-lines can be found in Bruhn-Suhr and Krumbholz (1990). Given a large-sized lot, a single sample $X_1, ..., X_n$, $(n > 3)$ with

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

an acceptable quality level p_1, a rejectable quality level p_2 and levels α and β of Type-I and Type-II error, respectively, Bruhn-Suhr and Krumbholz (1990) develop single sampling plans based on the essentially Maximum Likelihood (ML) estimator

$$p^* = p(\overline{X}, S) = \Phi\left(\frac{L - \overline{X}}{S}\right) + \Phi\left(\frac{\overline{X} - U}{S}\right). \tag{2}$$

The lot is accepted within the single plan $(n,\ k)$, if $p^* \leq k$.
With the help of the operation characteristic (OC) of single sampling plans, Vangjeli (2011) develops AM-double sampling plans λ_1^* based on the independent random variables p_1^* and p_2^*, which relate to the first and second samples, respectively. Given p_1, p_2, α, and β, the AM-double sampling plan fulfills the classical two-points-condition on the OC and features the lowest maximum of the average sample number (ASN). λ_1^* is computed in a similar fashion to the corresponding single sampling plan $(n,\ k)$ by using its one-sided approximation AM-double sampling plan $\widetilde{\lambda}_1$, which is based on information obtained only from the second sample in the second stage. A double sampling plan consisting of two independent consecutive samples needs a larger sample size to fulfill the classical two-points-condition on its OC than the corresponding double sampling plan defined by taking into account information from both samples in the second stage.
In this paper, we introduce the AM-double sampling plan λ_2^* based on the random variables p_1^* and p_p^*. Using the random variable p_p^*, which contains information from both samples in the second stage, the OC of an arbitrary double

sampling plan λ_2 becomes more complex than the OC of the corresponding double sampling plan λ_1. The probability for accepting the lot after the inspection of the first sample is analogously to λ_1 a single-sampling-plan-OC. Thus, in the next section some preliminaries regarding the single-sampling-plan-OC, as well as notation and definitions concerning the double sampling plan λ_2 are introduced. The increased complexity of λ_2-OC compared to λ_1-OC is found in the probability for accepting the lot after the inspection of the second sample. The derivation of this probability is described in Section 3. The AM-double sampling plan λ_2^* is computed analogously to λ_1^* by using the corresponding one-sided approximation AM-double sampling plan $\widetilde{\lambda}_2$. A comparison between λ_1^* and λ_2^* is presented in Section 4.

2. Preliminaries

Before introducing the notation and definitions for deriving the double-sampling-plan-OC, we first note a well-known issue from single sampling. Let

$$L_{(n,\ k)}(\mu, \sigma) = P(p^* \leq k) \tag{3}$$

denote the OC for the single plan $(n,\ k)$ and let g_r be the density function of the χ^2 distribution with r degrees of freedom.

Theorem 1: It holds that:

$$\begin{aligned} L_{(n,\ k)}(\mu, \sigma) = \int_0^B \Bigg\{ & \Phi\left(\frac{\sqrt{n}}{\sigma}\left(\mu\left(\sigma\sqrt{\frac{t}{n-1}}, k\right) - \mu\right)\right) \\ & -\Phi\left(\frac{\sqrt{n}}{\sigma}\left(\dot{\mu}\left(\sigma\sqrt{\frac{t}{n-1}}, k\right) - \mu\right)\right)\Bigg\} g_{n-1}(t)dt \end{aligned} \tag{4}$$

with

$$B = \frac{(n-1)(L-U)^2}{4\sigma^2\left(\Phi^{-1}\left(\frac{k}{2}\right)\right)^2} \quad \text{and} \quad \dot{\mu}(\sigma, p) = L + U - \mu(\sigma, p).$$

For the proof of Theorem 1, Bruhn-Suhr and Krumbholz (1990) use the fact that for a given $\mathring{p}$ $(0 < \mathring{p} < 1)$ and $\mathring{\sigma} > 0$,

$$M(\mathring{\sigma}, \mathring{p}) := \{\mu \in \mathbb{R} \mid p(\mathring{\sigma}, \mu) \leq \mathring{p}\} \tag{5}$$

is equivalent to

$$M(\mathring{\sigma}, \mathring{p}) = \begin{cases} [\dot{\mu}(\mathring{\sigma}, \mathring{p}), \mu(\mathring{\sigma}, \mathring{p})] & \text{if} \quad \mathring{\sigma} \leq \sigma_0(\mathring{p}) \\ \emptyset & \text{otherwise,} \end{cases} \tag{6}$$

with

$$\sigma_0(\mathring{p}) = \frac{L - U}{2\Phi^{-1}\left(\frac{\mathring{p}}{2}\right)}. \quad \text{(See Figure 1)} \tag{7}$$

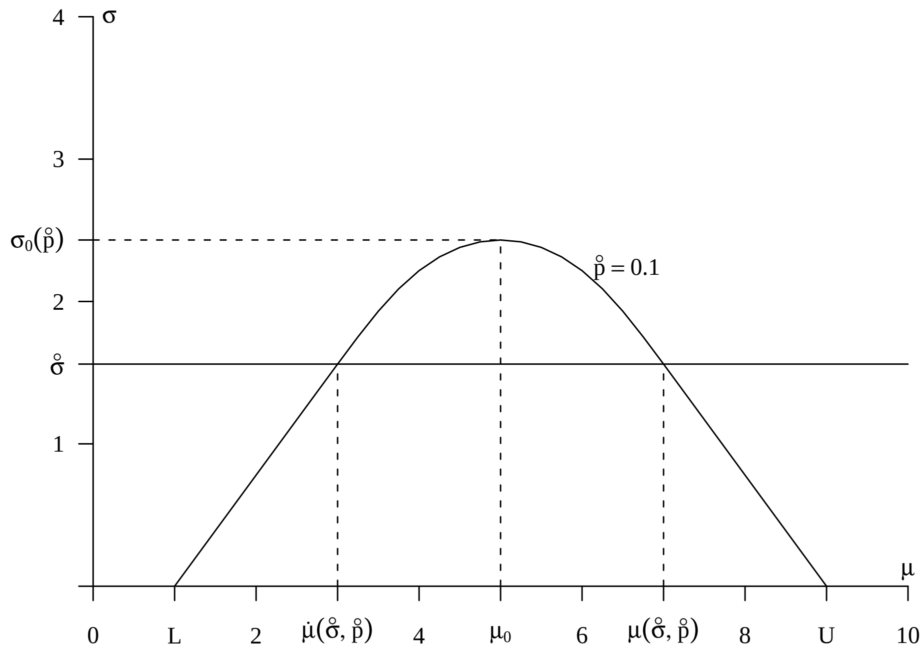

Figure 1: Iso-p-line for $\mathring{p} = 0.1$ with $\mu_0 = 5$, $\sigma_0 = 2.431827$ and $M(\mathring{\sigma}, \mathring{p})$ for $\mathring{\sigma} = 1.560192$

Now, we turn our attention to the double sampling plan λ_2. Let $X_1, ..., X_{n_1}$ be the first and $X_{n_1+1}, ..., X_{n_1+n_2}$ the second sample on X. Then, define the following notation:

$$\overline{X}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \tag{8}$$

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \overline{X}_1)^2 = \frac{1}{n_1 - 1} \left(\sum_{i=1}^{n_1} X_i^2 - n_1 \overline{X}_1^2 \right), \tag{9}$$

$$\overline{X}_2 = \frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} X_i, \tag{10}$$

$$\overline{\overline{X}} = \frac{1}{n_1 + n_2} \sum_{i=1}^{n_1+n_2} X_i = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2}, \tag{11}$$

$$S^2 = \frac{1}{n_1 + n_2 - 1} \sum_{i=1}^{n_1+n_2} (X_i - \overline{\overline{X}})^2. \tag{12}$$

Definition 1: The double sampling plan by variables $\lambda_2 = \begin{pmatrix} n_1 & k_1 & k_2 \\ n_2 & k_3 & \end{pmatrix}$ with $n_1, n_2 \in \mathbb{N}$; $n_1, n_2 \geq 2$; $k_1, k_2, k_3 \in \mathbb{R}^+$; $k_1 \leq k_2$, is defined as follows:

(i) Observe a first sample of size n_1 and compute $p_1^* = p(\overline{X}_1, S_1)$.

If $p_1^* \leq k_1$, accept the lot.
If $p_1^* > k_2$, reject the lot.
If $k_1 < p_1^* \leq k_2$, go to (ii).

(ii) Observe a second sample of size n_2 and compute $p_p^* = p(\overline{\overline{X}}, S)$.

If $p_p^* \leq k_3$, accept the lot.
If $p_p^* > k_3$, reject the lot.

The λ_2-OC is given by

$$L_{\lambda_2}(\mu, \sigma) = P_{(\mu,\sigma)}(A_1) + P_{(\mu,\sigma)}(A_2) \tag{13}$$

with

$$A_1 = \{p_1^* \leq k_1\},\ A_2 = \{p_p^* \leq k_3,\ k_1 < p_1^* \leq k_2\}. \tag{14}$$

From (3), (4) and (14) it follows that

$$P_{(\mu,\sigma)}(A_1) = L_{(n_1,\ k_1)}(\mu,\sigma). \tag{15}$$

Since $P(A_2) := P_{(\mu,\sigma)}(A_2)$ is more complex, we describe how to determine it in the next section. The λ_2-ASN is given by

$$N_{\lambda_2}(\mu,\sigma) = n_1 + n_2 P_{(\mu,\sigma)}(k_1 < p_1^* \leq k_2) \tag{16}$$

with

$$P_{(\mu,\sigma)}(k_1 < p_1^* \leq k_2) = L_{(n_1,k_2)}(\mu,\sigma) - L_{(n_1,k_1)}(\mu,\sigma).$$

Remark 1: The following analogies between λ_1 and λ_2 hold:

(i) $L_{\lambda_2}(\mu,\sigma)$ and $N_{\lambda_2}(\mu,\sigma)$ are not unique functions in p, but bands.

(ii) Let the symbol * indicate the AM-double sampling plan. Denoting ϕ_1^* as the one-sided AM-approximation for λ_1^*, Vangjeli (2011) shows that there are nonessential differences between $N_{max}(\lambda_1^*)$ and $N_{max}(\phi_1^*)$ [1].

3. The $P(A_2)$

Let

$$P(A_2^u) := P_{(\mu,\sigma)}(A_2^u) = P_{(\mu,\sigma)}(p_p^* \leq k_3,\ p_1^* \leq k_2) \tag{17}$$

and

$$P(A_2^l) := P_{(\mu,\sigma)}(A_2^l) = P_{(\mu,\sigma)}(p_p^* \leq k_3,\ p_1^* \leq k_1). \tag{18}$$

The probability

$$P(A_2) := P_{(\mu,\sigma)}(A_2) = P_{(\mu,\sigma)}(p_p^* \leq k_3,\ k_1 < p_1^* \leq k_2) \tag{19}$$

can be written as

$$P(A_2) = P(A_2^u) - P(A_2^l). \tag{20}$$

[1] The examples given in the fourth section confirm this fact for $N_{max}(\lambda_2^*)$ and $N_{max}(\phi_2^*)$

For $i = 1, 2$, let

$$Y_i := \sqrt{n_i}\,\frac{\overline{X}_i - \mu}{\sigma} \;\sim N(0,1) \tag{21}$$

and

$$W_i := \frac{n_i - 1}{\sigma^2}\,S_i^2 \;\sim \chi^2_{n_i-1}. \tag{22}$$

KRUMBHOLZ and ROHR (2006) have shown that the following holds:

$$S = \frac{\sigma\,\sqrt{(n_1+n_2)\,(W_1+W_2) + (\sqrt{n_2}\,Y_1 - \sqrt{n_1}\,Y_2)^2}}{\sqrt{(n_1+n_2-1)(n_1+n_2)}}. \tag{23}$$

Along with (21), it can be shown that

$$\overline{\overline{X}} = \frac{\sqrt{n_1}\,\sigma(Y_1 + \sqrt{n_1}\,\frac{\mu}{\sigma}) + \sqrt{n_2}\,\sigma(Y_2 + \sqrt{n_2}\,\frac{\mu}{\sigma})}{n_1+n_2}. \tag{24}$$

Due to total probability decomposition and the independence of $\overline{\overline{X}}$ and S^2, $P(A_2^u)$ can be written as:

$$P(A_2^u) = \int_0^\infty \Bigg(\int_{-\infty}^\infty \Bigg(\int_{-\infty}^\infty \Bigg(\int_0^\infty P(A_2^u | W_1 = w_1,\, Y_1 = y_1,\, Y_2 = y_2,\, W_2 = w_2) \times$$
$$\times\, g_{n_2-1}(w_2)\,dw_2\Bigg)\,\Phi'\,(y_2)\,dy_2\Bigg)\,\Phi'\,(y_1)\,dy_1\Bigg)\,g_{n_1-1}(w_1)\,dw_1. \tag{25}$$

It holds that:

$$P(A_2^u | W_1 = w_1,\, Y_1 = y_1,\, Y_2 = y_2,\, W_2 = w_2) =$$
$$= P(p(\overline{\overline{X}}, S) \le k_3,\; p(\overline{X}_1, S_1) \le k_2). \tag{26}$$

From (6) and (24), for $S < \sigma_0(k_3)$, we get:

$$p(\overline{\overline{X}}, S) \le k_3 \;\;\Leftrightarrow\;\; \dot{\mu}(S,k_3) \le \overline{\overline{X}} \le \mu(S,k_3) \;\;\Leftrightarrow\;\; C_1 \le Y_2 \le C_2, \tag{27}$$

where

$$C_1 = \frac{(n_1+n_2)\,\dot{\mu}(S,k_3) - (\sigma\sqrt{n_1}\,Y_1 + (n_1+n_2)\,\mu)}{\sigma\,\sqrt{n_2}} \tag{28}$$

and

$$C_2 = \frac{(n_1+n_2)\,\mu(S,k_3) - (\sigma\sqrt{n_1}\,Y_1 + (n_1+n_2)\,\mu)}{\sigma\,\sqrt{n_2}}. \tag{29}$$

From (23) and $S < \sigma_0(k_3)$, it follows that

$$W_2 \leq D, \tag{30}$$

where

$$D = \frac{(n_1+n_2)(n_1+n_2-1)\left(\frac{\sigma_0(k_3)}{\sigma}\right)^2 - (\sqrt{n_2}\,Y_1 - \sqrt{n_1}\,Y_2)^2}{n_1+n_2} - W_1.$$

Similarly, from $p(\overline{X}_1, S_1) \leq k_2$, we get:

$$E_1 \leq Y_1 \leq E_2 \tag{31}$$

with

$$E_1 = \frac{\sqrt{n_1}}{\sigma}\left(\dot{\mu}\left(\sigma\sqrt{\frac{W_1}{n_1-1}}, k_2\right) - \mu\right),$$

$$E_2 = \frac{\sqrt{n_1}}{\sigma}\left(\mu\left(\sigma\sqrt{\frac{W_1}{n_1-1}}, k_2\right) - \mu\right)$$

and

$$W_1 \leq F = \left(\frac{\sigma_0(k_2)}{\sigma}\right)^2 (n_1-1). \tag{32}$$

Setting $W_1 = w_1$, $Y_1 = y_1$, $Y_2 = y_2$, $W_2 = w_2$, $S = S(w_1,\, y_1,\, y_2,\, w_2)$, $C_1 = C_1(w_1,\, y_1,\, y_2,\, w_2)$, $C_2 = C_2(w_1,\, y_1,\, y_2,\, w_2)$, $D = D(w_1,\, y_1,\, y_2)$, $E_1 = E_1(w_1)$ and $E_2 = E_2(w_1)$, $P(A_2^u)$ can be written as:

$$P(A_2^u) = \int_0^F \Bigg(\int_{E_1(w_1)}^{E_2(w_1)} \Bigg(\int_{-\infty}^{\infty} \Bigg(\int_0^{D(w_1,y_1,y_2)} H(w_1,y_1,y_2,w_2) \times \times g_{n_2-1}(w_2)\,dw_2\Bigg)\,\Phi'\,(y_2)\,dy_2\Bigg)\,\Phi'\,(y_1)\,dy_1\Bigg)\,g_{n_1-1}(w_1)\,dw_1, \tag{33}$$

with

$$H(w_1, y_1, y_2, w_2) = \Phi(C_2(w_1, y_1, y_2, w_2)) - \Phi(C_1(w_1, y_1, y_2, w_2)).$$

$P(A_2^l)$ is obtained by substituting k_1 for k_2 in $P(A_2^u)$. Thus, we can state:

Theorem 2: It holds that:

$$L_{\lambda_2^*}(\mu, \sigma) = L_{(n_1,\ k_1)}(\mu, \sigma) + P(A_2^u) - P(A_2^l). \tag{34}$$

4. The computation of the AM-double sampling plans

For a given p_1, p_2, α and β, the plan λ_2^* is computed in a similar way as λ_1^*. We use the one-sided approximation $\widetilde{\lambda}_2 = \begin{pmatrix} n_1 & \tilde{k}_1 & \tilde{k}_2 \\ n_2 & \tilde{k}_3 & \end{pmatrix}$ with

$$\tilde{k}_1 = \Phi\left(\frac{l_1}{\sqrt{n_1}}\right),\ \tilde{k}_2 = \Phi\left(\frac{l_2}{\sqrt{n_1}}\right),\ \tilde{k}_3 = \Phi\left(\frac{l_3}{\sqrt{n_1 + n_2}}\right),$$

where $\phi_2^* = \begin{pmatrix} n_1 & l_1 & l_2 \\ n_2 & l_3 & \end{pmatrix}$ denotes the AM-double sampling plan in case of an upper tolerance limit U (cf. Krumbholz and Rohr (2009)). ϕ_2^* is determined by

$$\begin{aligned} &\text{(i)} \quad L_{\phi_2}(p_1) \geq 1 - \alpha \\ &\text{(ii)} \quad L_{\phi_2}(p_2) \leq \beta \\ &\text{(iii)} \quad N_{max}(\phi_2^*) = \min_{\phi_2 \in Z} N_{max}(\phi_2), \end{aligned} \tag{35}$$

where Z is the set of all double sampling plans ϕ_2 fulfilling (35)(i) and (ii). The AM-double sampling plan λ_2^* is given

$$\begin{aligned} &\text{(i)} \quad \min_{0<\sigma\leq\sigma_0(p)} L_{\lambda_2^*}(\sigma; p_1) \geq 1 - \alpha \\ &\text{(ii)} \quad \max_{0<\sigma\leq\sigma_0(p)} L_{\lambda_2^*}(\sigma; p_2) \leq \beta \\ &\text{(iii)} \quad N_{max}(\phi_2^*) = \min_{\phi_2 \in Z} N_{max}(\phi_2). \end{aligned} \tag{36}$$

Example 1

For $L = 1,\ U = 9,\ p_1 = 0.01,\ p_2 = 0.06,\ \alpha = \beta = 0.1$, we get:

(i) $(n,\ k) = (36,\ 0.02645943143)$ and $\alpha^* = 0.082,\ \beta^* = 0.1$,

(ii) $\lambda_1^* = \begin{pmatrix} 26 & 0.017577 & 0.035291 \\ 20 & 0.029275 & \end{pmatrix}$ and $N_{max}(\lambda_1^*) = 32.75439$.

α^{**}	β^{**}	$\widetilde{\lambda}_2$	$N_{max}(\phi_2^*)$	$\min_\sigma L_{\widetilde{\lambda}_2}(\sigma; p_1)$	$\max_\sigma L_{\widetilde{\lambda}_2}(\sigma; p_2)$
0.082	0.1	23 0.013909 0.038143 17 0.026289	30.45689	0.8930783818	0.0970618822
0.080	0.1	23 0.013597 0.038833 17 0.026400	30.72159	0.8955304122	0.0969993958
0.078	0.1	23 0.013993 0.038763 18 0.026496	30.99066	0.8978607677	0.0971362848
0.077	0.1	23 0.013838 0.039100 18 0.026558	31.12727	0.8990880486	0.0971046378
0.076	0.1	23 0.013681 0.039455 18 0.026617	31.26779	0.9003201617	0.0970742118

where $\lambda_2^* = \begin{pmatrix} 23 & 0.013681 & 0.039455 \\ 18 & 0.026617 & \end{pmatrix}$ with $N_{max}(\lambda_2^*) = 31.26778533$.

Example 2

For $L = 1,\ U = 9,\ p_1 = 0.01,\ p_2 = 0.03,\ \alpha = \beta = 0.1$, we get:

(i) $(n,\ k) = (115,\ 0.0178762881)$ and $\alpha^* = 0.085,\ \beta^* = 0.1$,

(ii) $\lambda_1^* = \begin{pmatrix} 81 & 0.014029 & 0.021742 \\ 66 & 0.018537 & \end{pmatrix}$ and $N_{max}(\lambda_1^*) = 103.5432$.

α^{**}	β^{**}	$\widetilde{\lambda}_2$	$N_{max}(\phi_2^*)$	$\min_\sigma L_{\widetilde{\lambda}_2}(\sigma;p_1)$	$\max_\sigma L_{\widetilde{\lambda}_2}(\sigma;p_2)$
0.085	0.1	72 0.012337 0.023495 58 0.017830	98.51959	0.8979623972	0.0991345737
0.084	0.1	72 0.012364 0.023535 59 0.017851	98.97047	0.8990715849	0.0991520345
0.083	0.1	72 0.012385 0.023569 60 0.017875	99.43030	0.9001786758	0.0991672779

where $\lambda_2^* = \begin{pmatrix} 72 & 0.012385 & 0.023569 \\ 60 & 0.017875 & \end{pmatrix}$ with $N_{max}(\lambda_2^*) = 99.43020285$.

Remark 2: Numerical investigations indicate:

(i) The AM-double sampling plan λ_2^* is more powerful than the AM-double sampling plan λ_1^* as it appears that

$$N_{max}(\lambda_2^*) < N_{max}(\lambda_1^*).$$

(ii) Let $\widehat{\lambda}_1$ denote the AM-double sampling plan based on the MVU estimators $\hat{p}_1$ and $\hat{p}_2$ of $p(\mu,\sigma)$. $\hat{p}_1$ and $\hat{p}_2$ are superior over p_1^* and p_2^*, respectively, so that

$$N_{max}(\widehat{\lambda}_1) < N_{max}(\lambda_1^*).$$

For some constellations, it could further be shown that

$$N_{max}(\lambda_2^*) < N_{max}(\widehat{\lambda}_1) < N_{max}(\lambda_1^*) \quad \text{(See Figure 2)}.$$

(iii) The lowest N_{max} among the AM-double sampling plans for a normally distributed quality characteristic with two-sided specification limits and unknown σ would feature the plan $\widehat{\lambda}_2$ based on the MVU estimators $\hat{p}_1$ and $\hat{p}_p$ of $p(\mu,\sigma)$, provided that a formula for determining the $\widehat{\lambda}_2$-OC would be found.

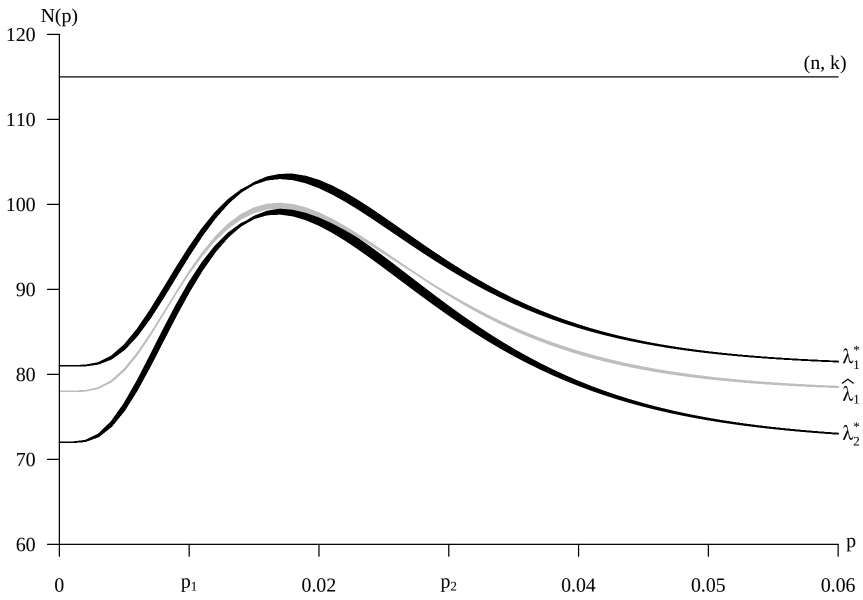

Figure 2: ASN bands for λ_1^*, $\widehat{\lambda}_1$ and λ_2^* defined by $p_1 = 0.01$, $p_2 = 0.03$ and $\alpha = \beta = 0.1$

References

Bruhn-Suhr, M., Krumbholz, W. (1990). A new variables sampling plan for normally distributed lots with unknown standard deviation and double specification limits. Statistical Papers 31, 195-207.

Krumbholz, W., Rohr, A. (2006). The operating characteristic of double sampling plans by variables when the standard deviation is unknown. Allgemeines Statistisches Archiv 90, 233-251.

Krumbholz, W., Rohr, A. (2009). Double ASN Minimax sampling plans by variables when the standard deviation is unknown. Advances in Statistical Analysis 93, 281-294.

Vangjeli, E. (2011). ASN-Minimax double sampling plans by variables for two-sided specification limits when σ is unknown.
http://arxiv.org/PS_cache/arxiv/pdf/1103/1103.4801v4.pdf.