A new one-sided variable inspection plan for continuous distribution functions

Wolfgang Kössler

Summary

The ordinary variable inspection plans rely on the normality of the underlying populations. However, often this assumption is vague or even not satisfied.

A new variable inspection plan is constructed that can be applied for continuous distributions with medium to long tails and that requires relatively moderate sample sizes only.

The main idea is that nonconforming items occur in the tails of the distribution. The tails are approximated by a generalized Pareto distribution and the fraction defective is estimated by the Maximum-Likelihood method.

Keywords: Attribute sampling plan, Extreme value distribution, Generalized Pareto distribution, Maximum-Likelihood estimation, Threshold.

1 Introduction

We consider a lot of units each having a quality characteristic $\mathbf{X}$ with a continuous cummulative distribution function (c.d.f.) F. Given a sample $X_1, ..., X_n$ a decision is to be made whether the lot is to be accepted or not. Assuming that there is only an upper specification limit U the fraction defective p of the lot is defined by

$$p = P(\mathbf{X} > U) = 1 - F(U).$$

If the upper tail of the c.d.f. F is not too small, reasonable estimates $\hat{p}$ of p based on the sample can be constructed. Our variable inspection plan is then defined as follows: If $\hat{p} \leq c, c \in (0,1)$, the lot will be accepted, else it will be rejected. The operating characteristic (OC)

$$L^{n,c}(p) := P_p(\hat{p} \leq c), \quad 0 < p < 1,$$

of the inspection plan constructed in this paper depends on the the underlying c.d.f. but not on the location and scale parameters.

Variable inspection plans are computed by minimizing the sample size n while meeting the *2-point conditions*

$$L^{n,c}(p_1) \geq 1 - \alpha \qquad \text{and} \tag{1}$$

$$L^{n,c}(p_2) \leq \beta \tag{2}$$

$(0 < p_1 < p_2 < 1, \quad 0 < \beta < 1 - \alpha)$.

For a normally distributed $\mathbf{X}$ the Lieberman-Resnikoff plan (LR-plan, cf. Resnikoff (1952), Lieberman and Resnikoff (1955)) can be applied. In the case of double specification limits, exact Maximum-Likelihood plans (ML-plans) are developed by Bruhn-Suhr and Krumbholz (1990). The OC of the ML-plan and the LR-plan are given by Bruhn-Suhr and Krumbholz (1990, 1991). Note that in our case of only one specification limit the LR-plan and the ML-plan are equivalent.

However, these inspection plans are very sensitive with respect to deviations from the normal distribution assumption (cf. e.g. Masuda (1978), Schneider and Wilrich (1981), Kössler and Lenz (1995, 1997)).

This problem gives rise to the question what should be done in the case of a non-normal or even unknown c.d.f.? One way is to perform sampling inspection by attributes, where an item i is considered to be nonconforming if and only if $X_i > U$. The OC is easy to compute (cf. Uhlmann (1982), ch.4.2.1). However, this method requires relatively large sample sizes.

In this paper we construct inspection plans that do not rely on the normality assumption and that require relatively moderate sample sizes only.

The main idea is that nonconforming items i occur in the tails of the underlying c.d.f., namely $X_i > U = F^{-1}(1-p)$ with the (unknown) fraction defective p. Additionally, items X_i with $X_i \approx U, X_i < U$, can be considered suspicious. They also should be considered in inspection plans.

The tails of a continuous c.d.f. can be approximated by a generalized Pareto distribution (GPD) as explained in Section 2. The approximation error is analysed in Section 3. In Section 4, estimates of the parameters of the GPD are used to obtain estimates of the fraction defective. In this paper we use Maximum-Likelihood estimators.

In Section 5 the asymptotic normality of these estimators is used to compute inspection plans that meet the conditions (1) and (2) at least approximately.

Simulation studies performed in Section 6 show that this method works well even for relatively small sample sizes. In Section 7 the new sampling plan is compared with corresponding attribute sampling plans.

2 Some preliminaries from the extreme value distribution theory

At first, assume that F is known. Let $X_{(i)}$ be the i-th order statistic of the sample $X_1, \ldots, X_n$ of size n. There are only 3 possible types of limiting c.d.f.s. for the normalized maximum $(X_{(n)} - b_n)/a_n$. These, along with choices for the centering and normalizing constants a_n and b_n are given as follows (take $\gamma > 0$, I denotes the indicator function and F^{-1} the quantile function):

$$
\begin{aligned}
G_{1,\gamma}(x) &:= exp(-x^{-\gamma}) \cdot I(x > 0), \\
&\qquad b_n = 0, \quad a_n = F^{-1}(1 - \frac{1}{n}), \\
G_{2,\gamma}(x) &:= exp(-(-x)^{\gamma}) \cdot I(x < 0) + I(x \geq 0), \\
&\qquad b_n = \sup\{x : F(x) < 1\} < \infty, \quad a_n = b_n - F^{-1}(1 - \frac{1}{n}), \\
G_{3,\gamma}(x) &:= exp(-exp(-x)) =: G_3(x), \\
&\qquad b_n = F^{-1}(1 - \frac{1}{n}), \quad a_n = h(b_n), \\
&\quad h(u) = \int_u^{\infty} (1 - F(v))dv/(1 - F(u)),
\end{aligned}
$$

when the last integral exists (cf. Reiss (1989), Ch. 5).

Remarks:

1. The constants a_n and b_n can be replaced by other constants a'_n and b'_n with $\frac{a'_n}{a_n} \to 1$ and $\frac{b'_n - b_n}{a_n} \to 0$.

2. The c.d.f. F is said to be in the domain of attraction of $G_{i,\gamma}$ if the limiting c.d.f. of the maximum $X_{(n)}$ is $G_{i,\gamma}$, $i = 1, 2, 3$. The three types of domains of attraction are denoted by Φ_γ, Ψ_γ and Λ, respectively. The c.d.f.s $G_{1,\gamma}$, $G_{2,\gamma}$ and $G_3 = G_{3,\gamma}$ are well-known as the Fréchet, “negative” Weibull and Gumbel c.d.f.

Definition: The generalized Pareto c.d.f. (GPD) is defined by

$$GPD(y; \sigma, k) := \begin{cases} 1 - (1 - \frac{ky}{\sigma})^{\frac{1}{k}} & \text{if } k \neq 0 \\ 1 - e^{-\frac{y}{\sigma}} & \text{if } k = 0, \end{cases}$$

where $\sigma > 0$. The range of y is given by $0 < y < \infty$ if $k \leq 0$ and $0 < y < \frac{\sigma}{k}$ if $k > 0$ (cf. Smith (1987), p. 1175).

Let t be an arbitrary real value of the support of the c.d.f. F and denote by $x_o := \sup\{x : F(x) < 1\}$ the upper endpoint of F, where $x_o = \infty$ is admissable.

The conditional c.d.f. $F_t(y)$ of $\mathbf{X} - t$ conditioned under $\mathbf{X} > t$ is given by

$$F_t(y) = \frac{F(t + y) - F(t)}{1 - F(t)}, \tag{3}$$

where $0 < y < x_o - t$. If $t \to x_o$ this distribution converges uniformly to a GPD with certain parameters σ and k as was shown by Pickands (1975), Theorem 7, i.e.

$$\lim_{t \to x_o} \|F_t(y) - GPD(y; \sigma(t), k)\|_\infty = 0.$$

To obtain the parameters k and σ we follow the arguments of Falk (1987). Since $F^n(x)$ is the c.d.f. of $X_{(n)}$ we have for $n \to \infty$:

$$\begin{array}{rcll} F^n(a_n x + b_n) & \to & G(x) & \text{iff} \\ n \log F(a_n x + b_n) & \to & \log G(x) & \text{iff} \\ n(1 - F(a_n x + b_n)) & \to & -\log G(x) & \text{iff} \\ \dfrac{1 - F(a_n x + b_n)}{1 - (1 + \log(G^{\frac{1}{n}}(x)))} & \to & 1. & \end{array}$$

Now

$$1 + \log G_{i,\gamma}(x) = \begin{cases} 1 - x^{-\gamma} & \text{if} \quad i = 1 \\ 1 - (-x)^\gamma & \text{if} \quad i = 2 \\ 1 - e^{-x} & \text{if} \quad i = 3 \end{cases} =: W_{i,\gamma}.$$

Hence,

$$1+\log G_{i,\gamma}^{\frac{1}{n}}(x)=\begin{cases} W_{1,\gamma}(n^{\frac{1}{\gamma}}x) & \text{if} \quad i=1 \\ W_{2,\gamma}(n^{\frac{1}{-\gamma}}x) & \text{if} \quad i=2 \\ W_{3,\gamma}(x+\log n) & \text{if} \quad i=3 \end{cases} \qquad =: W_{(n)}(x).$$

Given $a_n, b_n, t>0$, let $x_t=\frac{t-b_n}{a_n}$. Then

$$\begin{aligned} 1-F(t) &= 1-F(a_n x_t+b_n) \\ &\sim \begin{cases} (n^{\frac{1}{\gamma}}x_t)^{-\gamma} & \text{if} \quad i=1 \\ (-n^{-\frac{1}{\gamma}}x_t)^{\gamma} & \text{if} \quad i=2 \\ e^{-(x_t+\log n)} & \text{if} \quad i=3. \end{cases} \\ &= \frac{1}{n}\cdot\begin{cases} (\frac{t}{a_n})^{-\gamma} & \text{if} \quad i=1 \\ (-\frac{t-b_n}{a_n})^{\gamma} & \text{if} \quad i=2 \\ e^{-(\frac{t-b_n}{a_n})} & \text{if} \quad i=3, \end{cases} \end{aligned}$$

where the symbol $\sim$ denotes asymptotic equivalence. Therefore,

$$\begin{aligned} F_t(y) &= 1-\frac{1-F(t+y)}{1-F(t)} \\ &\sim 1-\begin{cases} (1+\frac{y}{t})^{-\gamma} & \text{if} \quad i=1 \\ (1+\frac{y}{t-b_n})^{\gamma} & \text{if} \quad i=2 \\ e^{-\frac{y}{a_n}} & \text{if} \quad i=3 \end{cases} \\ &= GPD(y;\sigma,k), \end{aligned}$$

where

$$k = \begin{cases} -\frac{1}{\gamma} & \text{if} \quad i=1 \\ \frac{1}{\gamma} & \text{if} \quad i=2 \\ 0 & \text{if} \quad i=3, \end{cases} \tag{4}$$

$$\sigma = \begin{cases} -kt & \text{if} \quad i=1 \\ -k(t-x_o) & \text{if} \quad i=2 \\ h(t) & \text{if} \quad i=3. \end{cases} \tag{5}$$

If $i=2$ then $b_n=x_o<\infty$. If $i=3$ then n is chosen such that $b_n\le t<b_{n+1}$ and $a_n=h(t)$, where the function h is defined as above.

If F is a GPD with parameters τ and $k, \tau>kt$, then F_t is also a GPD, now with parameters $\sigma(t)=\tau-kt$ and k. Special cases are the Pareto, exponential, triangle and uniform.

3 Approximation of the fraction defective p

Now we use the GPD to approximate the fraction defective p and analyse the approximation error. Let t be fixed, $t < U = F^{-1}(1-p)$; $\quad y = U - t$ and $\Delta_t(y) = F_t(y) - GPD(y;\sigma(t),k)$, where σ and k are defined by (5) and (4). Then we obtain from (3):

$$1 - p = F(U) = F(t) + F_t(y) \cdot (1 - F(t))$$

and hence,

$$p = (1 - F(t))(1 - F_t(y)) = (1 - F(t)) \cdot (1 - GPD(y;\sigma,k) + \Delta_t(y)).$$

If we use the GPD-approximation we estimate instead of p the approximation

$$p_F = (1 - F(t)) \cdot \begin{cases} (1 - \frac{ky}{\sigma})^{\frac{1}{k}} & \text{if} \quad k \neq 0 \\ e^{-\frac{y}{\sigma}} & \text{if} \quad k = 0, \end{cases}$$

with the (absolute) error

$$p - p_F = \Delta_t(y) \cdot (1 - F(t)).$$

If t is an $(1-q)$-quantile of the underlying c.d.f. F with $q > p$, then $p_F = q \cdot (1 - GPD(y;\sigma,k))$ and $p - p_F = \Delta_t(y) \cdot q$.

For some c.d.fs., e.g. for the Fréchet, i.e. $F(x) = e^{-x^{-\gamma}}, \quad \gamma > 0$, the relative error is small, $\frac{p-p_F}{p} \cdot 100\% \leq 10\%$ in almost all cases of parameters γ and fraction defectives $p, p \leq 0.2$. For some other distributions, e.g. for the normal, these approximations are worse (cf. Table 2).

But this is not the problem. We are more interested in an approximation for finite, fixed t. For fixed t the parameters σ and k given from the extreme value distribution theory may not be the best. This was already realized by DuMouchel (1983) who obtained "optimal" values for k as solutions from Maximum-Likelihood (ML) equations in the GPD-model. He choosed $t = F^{-1}(0.9)$ for a given c.d.f. F and constructed pseudodata by $x_j = F^{-1}(0.9000 + j \cdot 0.0001), \quad j = 0, \ldots, 999$.

We use this idea with some modifications and additions. We choose the smaller value of $t = F^{-1}(0.8)$, pseudodata $x_j = F^{-1}(0.8 + j \cdot 0.2/N), j = 0, \ldots, N-1$, where N is the size of the pseudosample. The solutions of the ML-equations are called *pseudo-ML-estimates*. They result in estimates $\hat{p}_F$ of p_F. The estimated relative approximation errors $\widehat{err}(p,F) = \frac{p-\hat{p}_F}{p} \cdot 100\%$ are also computed for some $p < 0.2$. For $N = 2000$ the results are presented in Tables 1 and 2.

Table 1: The parameters σ and k "estimated" by ML compared with that from the asymptotic theory.

	pseudo ML-estimates		asymptotic values	
c.d.f.	σ	k	σ	k
Pareto (1)	5.016	-0.993	5.000	-1.000
Pareto (1.5)	1.960	-0.659	1.949	-0.667
Pareto (2)	1.121	-0.493	1.118	-0.500
Pareto (5)	0.275	-0.196	0.276	-0.200
Cauchy	1.744	-0.948	1.376	-1.000
Fréchet(1)	5.029	-0.992	4.481	-1.000
Fréchet(1.5)	2.021	-0.647	1.812	-0.667
Fréchet(2)	1.173	-0.477	1.058	-0.500
Fréchet(5)	0.296	-0.169	0.270	-0.200
Exponential	1.005	0.008	1.000	0.000
Logistic	1.182	0.062	1.116	0.000
Normal	0.663	0.192	0.559	0.000

From the entries for the Pareto and for the exponential distribution in table 1 it can be observed that even for "ideal" data and large sample sizes the ML estimates of σ and k have a bias. But for $N = 2000$ this bias is small, except perhaps for the normal and Cauchy, resulting in small relative errors, $err(p, F) \leq 7\%$ for all distributions considered and all p with $p \in \{0.005, 0.01, 0.05, 0.1, 0.15\}$ (cf. Table 2). If $p \geq 0.05$ or if the c.d.f. is not normal or Cauchy, these errors are much smaller. If we choose $N = 20000$ or $N = 200$ the differences between the "optimal" and the asymptotic σ and k, will become smaller respectively larger.

This way the fraction defective is approximated by using the GPD. Furthermore, it is shown that for a suitably chosen, fixed t the approximation error is relatively small.

4 Estimation of the fraction defective p in the GPD-model

Let t be a threshold, $y = U - t$ and F_n the empirical c.d.f. of the sample. Now assume that F_t is $GPD(\cdot; \sigma, k)$ with unknown σ and k, $k < 0.5$, i.e. the unknown F has not too short tails. For comparison, if F is a triangle c.d.f. then $k = 0.5$. In the context of quality control all interesting distributions satisfy

Table 2: "Estimated" relative errors $\frac{p-\hat{p}_F}{p} \cdot 100\%$ in the approximation of F_t by GPD

	p				
c.d.f.	0.005	0.01	0.05	0.1	0.15
Pareto(1)	-1.578	-1.124	-0.208	0.023	0.052
Pareto(1.5)	-1.808	-1.178	-0.053	0.139	0.108
Pareto(2)	-2.420	-1.647	-0.271	0.015	0.050
Pareto(5)	-3.109	-2.266	-0.720	0.277	0.093
Cauchy	-5.541	-2.593	1.105	0.255	0.403
Fréchet(1)	-1.725	-1.188	0.169	0.032	0.039
Fréchet(1.5)	-2.393	-1.155	0.242	0.109	-0.024
Fréchet(2)	-2.845	-1.256	0.206	-0.044	-0.152
Fréchet(5)	-3.182	-0.944	0.291	-0.257	-0.346
Exponential	1.864	1.511	0.696	-0.348	-0.144
Logistic	-0.722	2.509	1.130	-0.794	-0.972
Normal	5.475	6.926	0.715	-1.652	-1.465

this condition.

Let $\hat{k}$ and $\hat{\sigma}$ be consistent estimators of k and σ, respectively, in the GPD-model. Then

$$\hat{p_o} = (1 - F_n(t)) \cdot \begin{cases} (1 - \frac{\hat{k}y}{\hat{\sigma}})^{\frac{1}{\hat{k}}} & \text{if} \quad \hat{k} \neq 0 \\ e^{-\frac{y}{\hat{\sigma}}} & \text{if} \quad \hat{k} = 0 \end{cases} \tag{6}$$

is a consistent estimator of p. Now specify

$$t = X_{(n-m)} \tag{7}$$

where $m := \lceil (n+1)q \rceil$ with some $q, \quad 0 < q < 1, \quad$ and let $\lceil A \rceil$ denote the smallest integer greater than or equal to A. Then

$$\hat{p} = q \cdot \begin{cases} (1 - \frac{\hat{k}y}{\hat{\sigma}})^{\frac{1}{\hat{k}}} & \text{if} \quad \hat{k} \neq 0 \\ e^{-\frac{y}{\hat{\sigma}}} & \text{if} \quad \hat{k} = 0 \end{cases} \tag{8}$$

is also a consistent estimate of p.

Let $y_j = X_{(n-m+j)} - t, \quad j = 1, \ldots, m$, $\mathbf{Y} = (y_1, \ldots, y_m)$ with t defined by (7).

4.1 The ML-estimation

The ML estimates $\hat{k}_{ML}$ and $\hat{\sigma}_{ML}$ of k and σ, respectively, in the GPD-model are given by $\text{argmax}_{\sigma,k} l(\mathbf{Y};\sigma,k)$, where

$$l(\mathbf{Y};\sigma,k) = \begin{cases} -m\log\sigma + (\frac{1}{k}-1)\sum_{j=1}^{m}\log(1-\frac{ky_j}{\sigma}) & \text{if} \quad k \neq 0 \\ -m\log\sigma - \sum_{j=1}^{m}\frac{y_j}{\sigma} & \text{if} \quad k = 0 \end{cases}$$

is the log-likelihood of $\mathbf{Y}$. Since the ML-estimates $\hat{k}_{ML}$ and $\hat{\sigma}_{ML}$ are consistent, $\hat{p}$ given by (8) is also consistent.

Note that the estimates $\hat{k}_{ML}$ and $\hat{\sigma}_{ML}$ are invariant with respect to changes of the location parameters of F.

Recall that $x_o = \sup\{x : F'(x) > 0\}$. Under certain conditions on the rates of convergence of $t \to x_o$, $U \to x_o$ if $n \to \infty$, the ML-estimate $\hat{p}_{ML}$ is asymptotically normally distributed with expectation zero:

$$\mathbf{Z} := \sqrt{m}\frac{\hat{p}_{ML}-p}{p} \to N(0,V). \tag{9}$$

The details on the rates of convergence are given in Smith (1987). Since they are more of theoretical interest they are omitted here.

To obtain a closed relation for V dependent on F we follow the arguments of Smith (1987), Ch. 8, 9.

At first, assume that $k \neq 0$, i.e. F is attracted by a Fréchet- or "negative" Weibull-c.d.f. Let $z, \quad z > 0$, be fixed and $p_m, q_m, \quad p_m \to 0, \quad q_m \to 0, \quad 0 < p_m < q_m$ be defined such that

$$z = \frac{F^{-1}(1-p_m)}{F^{-1}(1-q_m)}$$

for every $m, \quad m = 1, 2, \ldots$.

Denote the nominator and the denominator of the last fraction by U_m and T_m, respectively. With $v_m := U_m - T_m$ we obtain $z = 1 + \frac{v_m}{T_m} = 1 - \frac{kv_m}{\sigma_m}$, where σ_m is given by (5). Hence z is defined in the same way as in Smith (1987). The asymptotic variance $V = V_F$ for $p_m, q_m \to 0$ is then given by

$$V_F = 1 + \mathbf{c}^T\mathbf{S}\mathbf{c} \tag{10}$$

where

$$\mathbf{c}^T = \left(-\frac{1}{k}(\frac{1}{z}-1), \frac{\log z}{k^2} + \frac{1}{k^2}(\frac{1}{z}-1)\right) \tag{11}$$

and

$$\mathbf{S} = (1-k)\begin{pmatrix} 2 & 1 \\ 1 & 1-k \end{pmatrix}, \tag{12}$$

cf. SMITH (1987), Ch. 8.

Now assume that $k = 0 \quad (F \in \Lambda)$, i.e. F is attracted by a Gumbel c.d.f. Let $\zeta, \quad \zeta > 0$ be fixed, $\psi(T) = \frac{1-F(T)}{f(T)}$, where f is the density corresponding to F and $p_m, q_m, \quad p_m \to 0, \quad q_m \to 0, \quad 0 < p_m < q_m$ be defined such that

$$\zeta = \frac{F^{-1}(1-p_m) - F^{-1}(1-q_m)}{\psi(F^{-1}(1-q_m))}$$

for every $m = 1, 2 \ldots$.

Hence, $\zeta = \frac{U_m - T_m}{\psi(T_m)} = \frac{v_m}{\psi(T_m)}$ is also defined in the same way as in Smith (1987), Ch. 9. The asymptotic variance $V = V_F$ for $p_m, q_m \to 0$, is then given by

$$V_F = 1 + 2\zeta^2 - \zeta^3 + \frac{\zeta^4}{4}, \tag{13}$$

cf. SMITH (1987), Theorem 9.5. Following the derivation of SMITH (1987), Ch. 8, we see that the asymptotic variance is also represented by

$$V_{F,q_m} = \begin{cases} 1 - q_m + \mathbf{c}^T\mathbf{S}\mathbf{c} & \text{if} \quad k \neq 0 \\ 1 - q_m + 2\zeta^2 - \zeta^3 + \frac{\zeta^4}{4} & \text{if} \quad k = 0. \end{cases} \tag{14}$$

This form yields a better approximation to the variance of $\mathbf{Z}$.

The asymptotic variance depends on q_m, but essentially on the choice of z and ζ, respectively. Recall that $p = 1 - F(U)$ is the fraction nonconforming. For given $q, \quad q > p$, let T be the $(1-q)$-quantile of F, i.e. $q = 1 - F(T)$. Hence, we can define z and ζ by

$$z = \frac{U}{T} = \frac{F^{-1}(1-p)}{F^{-1}(1-q)} \tag{15}$$

and

$$\zeta = \frac{U - T}{\psi(T)} = \frac{F^{-1}(1-p) - F^{-1}(1-q)}{\psi(F^{-1}(1-q))}, \tag{16}$$

respectively.

In the following we use the representation $V_{F,q}$ given by (14), where $\mathbf{c}, \mathbf{S}, z$, and ζ are defined by (11), (12), (15) and (16), respectively. In Table 3 some of these variances are presented.

The entries on the left hand side of this table denote the c.d.f., eventually with some shape parameter included in parentheses. The abbreviation "neg.Weibull" denotes the "negative" Weibull c.d.f.

Table 3: The asymptotic variances $V_{F,q}$ of $\hat{p}_{ML}$ given by (14)

		p						
		0.01		0.05			0.1	
		q		q			q	
type	c.d.f.	0.2	0.1	0.3	0.2	0.1	0.3	0.2
Φ_γ	Pareto(1)	13.38	6.96	3.96	2.76	1.66	2.07	1.56
	Pareto(1.5)	13.43	6.81	3.92	2.80	1.66	2.12	1.56
	Pareto(2)	13.33	6.70	3.91	2.82	1.66	2.14	1.56
	Pareto(5)	12.75	6.47	3.92	2.88	1.62	2.16	1.52
	Cauchy	15.18	7.20	5.84	3.13	1.70	2.81	1.72
	Fréchet(1)	14.66	7.30	4.62	2.98	1.70	2.30	1.64
	Fréchet(1.5)	14.82	7.15	4.54	3.01	1.70	2.34	1.65
	Fréchet(2)	14.76	7.03	4.51	3.03	1.70	2.37	1.65
	Fréchet(5)	14.17	6.77	4.46	3.08	1.67	2.41	1.62
Ψ_γ	neg.Weibull(2.2)	10.58	6.42	4.49	3.01	1.45	2.24	1.39
	neg.Weibull(2.5)	10.83	6.42	4.50	3.05	1.48	2.28	1.43
	neg.Weibull(3)	11.23	6.42	4.51	3.09	1.51	2.33	1.46
	neg.Weibull(5)	12.10	6.46	4.49	3.12	1.57	2.38	1.52
Λ	exponential	12.00	6.32	3.95	2.90	1.59	2.15	1.49
	logistic	7.92	5.58	3.01	2.57	1.55	1.84	1.41
	normal	5.13	4.28	2.59	2.31	1.50	1.71	1.36
	Weibull(1.2)	9.04	5.60	3.49	2.72	1.56	2.02	1.45
	Weibull(1.5)	7.11	5.02	3.12	2.56	1.53	1.90	1.42
	Weibull(2)	5.82	4.55	2.81	2.41	1.51	1.80	1.39
	Weibull(3)	4.91	4.16	2.56	2.28	1.49	1.70	1.35
	lognormal(0.3)	8.31	5.56	3.21	2.63	1.55	1.92	1.43
	lognormal(0.5)	13.08	6.94	3.83	2.90	1.60	2.09	1.48
	lognormal(0.8)	32.45	10.66	5.13	3.40	1.66	2.38	1.56
	lognormal(1.0)	67.71	15.34	6.66	3.83	1.71	2.63	1.62
	χ^2_4	9.97	5.93	3.59	2.82	1.59	2.05	1.48
	χ^2_6	9.03	5.63	3.45	2.77	1.59	2.02	1.48
	χ^2_8	8.45	5.43	3.36	2.74	1.58	2.00	1.47
	χ^2_{12}	7.75	5.16	3.26	2.69	1.57	1.97	1.46

From Table 3 we conclude that the relative differences between the variances are small if the fraction p/q is not too small (and perhaps if the c.d.f. is not lognormal). Therefore we construct inspection plans using the asymptotic relation (9) with the variances computed from a Pareto(1)- distribution.

To do this we have to choose the (theoretical) threshold T or, equivalently, the quantile q. Now the upper tolerance limit U is fixed and T has to be smaller than U. A problem is that the asymptotics is valid if $T \to x_o$. On the other hand, if m is large and the fraction defective p is small the asymptotic relation (9) can be used.

4.2 The choice of t

Choosing the practical threshold t we have to consider the following facts: On one hand the GPD-approximation works only if $t < U$. On the other hand if $t \geq U$ a decision is to be made whether the lot is to be accepted or not.

If t is an order statistic, say $t = X_{(r)}$, then we have

$$\begin{array}{lll} t \geq U & \text{if and only if} & F^{-1}(t) \geq F^{-1}(U) \\ & \text{if and only if} & U_{(r)} \geq 1 - p \\ & \text{if and only if} & p \geq 1 - U_{(r)} = 1 - s + s - U_{(r)}, \end{array}$$

where $U_{(r)}$ is the r-th order statistic of a random variable uniformly distributed on $(0,1)$. The distribution of $U_{(r)}$ can be reformulated using the Binomial or the $\mathbf{F}$-distribution. Let $Bi(\kappa; n, s)$ be the binomial c.d.f. with parameters n and s and $F_{\nu_1,\nu_2}(x)$ the c.d.f. of the $\mathbf{F}$-distribution with (ν_1, ν_2) degrees of freedom. Then we have

$$\begin{aligned} P(s - U_{(r)} \geq 0) &= P(U_{(r)} \leq s) = Bi(n - r; n, 1 - s) = \\ &= 1 - F_{\nu_1,\nu_2}\left(\frac{r}{n - r + 1} \cdot \frac{1 - s}{s}\right) = \\ &= F_{\nu_2,\nu_1}\left(\frac{n - r + 1}{r} \cdot \frac{s}{1 - s}\right), \end{aligned}$$

where $\nu_1 = 2(n - r + 1)$ and $\nu_2 = 2r$.

Now let q be associated with r by $r = \lfloor (n+1)(1-q) \rfloor$, where $\lfloor A \rfloor$ denotes the largest integer less than or equal to A. Choosing $s = 1 - p_2$, where p_2 is given from the 2-point condition (2), the relation $t \geq U$ implies $p \geq p_2$ with probability at least $F_{\nu_2,\nu_1}(\frac{q}{1-q} \cdot \frac{1-p_2}{p_2}) \approx F_{2n(1-q),2nq}(\frac{q}{1-q} \cdot \frac{1-p_2}{p_2})$. Consider this as a function of p_2 and q for different n. Looking at the shape of the function we see that for $q \geq p_2 + 0.1$ the probability of $p > p_2$ conditioned on $t > U$ is greater than 0.9.

Therefore, if $t \geq U$ then p is large with high probability and the lot should be rejected because of low quality.

This way, the choice of $t = X_{(r)}$, with $r = \lfloor (n+1)(1-q) \rfloor$ and $q = p_2 + 0.1$, is sufficiently motivated.

Note that, for $p_2 \to 0$, we have $q \to 0.1$, which seems to be sufficiently small for using the asymptotic theory.

5 Determination of a new sampling plan

In the last section we have established a relation between the sample size n and the threshold t. Given $q > 0$ define m by

$$m = n - \lfloor (n+1)(1-q) \rfloor, \tag{17}$$

i.e. for the estimation of the fraction defective only the m largest observations are used.

Since m is essential, the sampling plan is denoted by (n, m, c). An approximate OC of this sampling plan is given by the asymptotic distribution of $\hat{p}_{ML}$. At first, assume that the underlying c.d.f. is a Pareto(1). Then the variance $V(p)$ is given by $(14), (11)$ and (12) with $k = -1$.

To determine m and c meeting the two-point conditions (1) and (2) approximately we solve the system of equations

$$P_{p_1}(\hat{p} < c) \approx \Phi(\sqrt{m}\frac{c - p_1}{p_1\sqrt{V(p_1)}}) = 1 - \alpha \tag{18}$$

$$P_{p_2}(\hat{p} < c) \approx \Phi(\sqrt{m}\frac{c - p_2}{p_2\sqrt{V(p_2)}}) = \beta \tag{19}$$

The (unique) solution (m', c) of this system of equations is given by

$$m' = \frac{1}{(p_1 - p_2)^2}(p_2\sqrt{V(p_2)}\Phi^{-1}(\beta) - p_1\sqrt{V(p_1)}\Phi^{-1}(1-\alpha))^2 \tag{20}$$

$$c = p_1 + \Phi^{-1}(1-\alpha)\frac{p_1\sqrt{V(p_1)}}{\sqrt{m'}} \tag{21}$$

Since m' is generally not an integer, all the pairs (m,c) with $m = \lceil m' \rceil$ and $c \in [c_l, c_r]$, where

$$c_l = p_1 + \Phi^{-1}(1-\alpha)\frac{p_1\sqrt{V(p_1)}}{\sqrt{m}} \tag{22}$$

$$c_r = p_2 + \Phi^{-1}(\beta)\frac{p_2\sqrt{V(p_2)}}{\sqrt{m}} \tag{23}$$

meet the two-point conditions approximately.

Given m and q the sample size n is determined by the solution of (17). which is given by $n = \lfloor \frac{m}{q} \rfloor$.

In such a way a new sampling plan (n, m, c) is obtained. It is given by

$$n = \lfloor \frac{m}{q} \rfloor, \qquad m = \lceil m' \rceil \qquad \text{where} \qquad q = p_2 + 0.1$$

and m' and c are given by (20) and (21) . Note that $V(p)$ is also dependent on q. The approximate OC is given by

$$L_{ML}^{n,m,c}(p) = \Phi(\sqrt{m}\frac{c-p}{p\sqrt{V(p)}}) \tag{24}$$

Since the underlying c.d.f. is unknown, we use the variances $V(p)$ given from the Pareto(1) for all c.d.f.s.

In Table 4 the sampling plans (n, m, c) for 10 different 2-point conditions are presented. For comparison the corresponding sample sizes n_V and n_A of the LR-variable sampling plan and the attribute sampling plan, respectively, are given in the last two columns of Table 4. (For the definition of c_1, see the next section.)

The examples 1,2,4,5,7 and 8 are adopted from Resnikoff (1952), Tables 3, 5 and 7. They result in "round" sample sizes n_V for the LR-plan. The computation of the sample sizes n_A for the attribute sampling plan is described in Uhlmann (1982), Ch.5.2.2.

From Table 4 it can be seen that the sample sizes for the new plan are considerably less than that for the attribute sampling plan.

But how sensitive is the new plan with respect to deviations from the Pareto(1)? This question is discussed in the next section.

6 Simulation studies

The method described to obtain variable sampling plans is based on the asymptotic normality of the estimates $\hat{p}$ with the variance $V(p)$. The reference c.d.f. for computing $V(p)$ is the Pareto(1).

Table 4: Sample sizes for the new and for the ordinary variable sampling plans together with that for the attribute sampling plan.

Nr.	p_1	$1-\alpha$	p_2	β	n	m	c	c_1	n_V	n_A
1	0.0521	0.9500	0.1975	0.10	31	9	0.10845	0.1189	25	51
2	0.0634	0.9000	0.1975	0.10	34	10	0.11065	0.1204	25	50
3	0.0100	0.9000	0.0600	0.10	63	10	0.02398	0.0251	33	87
4	0.0100	0.9743	0.0592	0.10	82	13	0.02834	0.0294	50	110
5	0.0152	0.9000	0.0592	0.10	88	14	0.02956	0.0306	50	110
6	0.0100	0.9900	0.0600	0.10	88	14	0.03066	0.0317	61	130
7	0.0360	0.9500	0.0866	0.10	140	26	0.05806	0.0593	100	187
8	0.0406	0.9000	0.0866	0.10	145	27	0.05857	0.0598	100	187
9	0.0100	0.9900	0.0600	0.01	194	31	0.02398	0.0244	106	258
10	0.0100	0.9900	0.0300	0.10	362	47	0.02020	0.0204	205	469

To investigate whether the new sampling plan can be applied for moderate sample sizes as well as for c.d.f.s different from the Pareto(1), simulation studies are carried out. The OC is estimated and 0.95-confidence regions are constructed at the points p_1 and p_2. The ten examples from Table 4 are investigated. The simulation size is $M = 2000$. The following c.d.f.s are included in the simulation study: Pareto(1), Pareto(2), Cauchy, Fréchet(1), Fréchet(2), normal, logistic, exponential and triangle. These c.d.f.s are abbreviated by P1, P2, CA, FR1, FR2, N, L, E and Δ respectively.

To produce uniform random numbers the generator G2UI01 from the program package NUMATH2, cf. PP NUMATH (1984), is used. This random number generator is based on the algorithm $v_{i+1} = v_i \cdot 7^5 \mod (2^{31} - 1)$, $i = 1, 2, \ldots$, with normalizing on (0,1). The starting integer is generated by the actual computer time.

Samples from a population with c.d.f. F are produced by the transformation $x = F^{-1}(u), \quad u \in (0, 1)$. The computations are performed in FORTRAN on an IBM 386 computer.

Some approximate 0.95−confidence intervals $I_{\widehat{OC}} = (A_1, A_2)$ are given in Table 5.

First simulations show that the OC is, dependent on the sample size, more or less shifted to the left. The reasons for this behaviour may be bias and skewness in the estimation of p.

To compensate this property the acceptance number c is shifted to the right by $c_1 := c \cdot (1 + 3/n)$. The new sampling plans (n, c_1) are now considered in the simulation study. The Table 6 contains the estimated OC-values at the points p_1 and p_2 for the ten examples and the densities above.

For the Pareto(1) and the Fréchet(1) all the values p_1, p_2 are located in

Table 5: Approximate confidence intervals $I_{\widehat{OC}}$ for the OC dependent on $\widehat{OC}$

$\widehat{OC}$	A_1	A_2	$\widehat{OC}$	A_1	A_2
0.995	0.9919	0.9981	0.005	0.0019	0.0081
0.990	0.9856	0.9944	0.010	0.0056	0.0144
0.985	0.9797	0.9903	0.015	0.0097	0.0203
0.975	0.9682	0.9818	0.025	0.0182	0.0318
0.960	0.9514	0.9686	0.040	0.0314	0.0486
0.950	0.9404	0.9596	0.050	0.0404	0.0596
0.940	0.9296	0.9504	0.060	0.0496	0.0704
0.910	0.8975	0.9225	0.090	0.0775	0.1025
0.900	0.8869	0.9131	0.100	0.0869	0.1131
0.890	0.8763	0.9037	0.110	0.0963	0.1237

the corresponding confidence region. For the other densities there are some examples with somewhat too small $(1-\alpha)$-values, and very few examples with too large β-values. Large β-values occur only in cases of densities with very large (Cauchy) or very small (triangle) tails. For the normal we have some examples (6-10) with somewhat too small $(1-\alpha)$, whereas for the other densities there are few examples with small $(1-\alpha)$. Even for the Cauchy the OC-curve is not far from the two points $(p_1, 1-\alpha)$ and (p_2, β).

For illustration, Example 7 (cf. Table 4) is considered and three OC-curves are presented in Figure 1. The continuous line is the curve of the asymptotic OC $L_{ML}^{n,m,c}(p)$ of the new sampling plan if the density is Pareto(1), the dotted line is the worst case simulated OC-curve (here the worst case is attained for the normal) and the dashed line is the OC-curve of the corresponding attribute plan (n, c_A) (cf. Section 7).

Comparing the results of the robustness study with that obtained for the ML- or LR-variable sampling plans (cf. Kössler and Lenz (1995, 1997)), the new variable sampling plan seems to be very promising.

7 Comparison of the new variable sampling plans with attribute sampling plans

From the simulation study it can be seen that the two-point conditions are satisfied in most cases and if not, then the OC-values do not differ much from the nominal values. This is valid for all densities considered, although

Table 6: Estimated OC at the points p_1 and p_2 for the new sampling plan

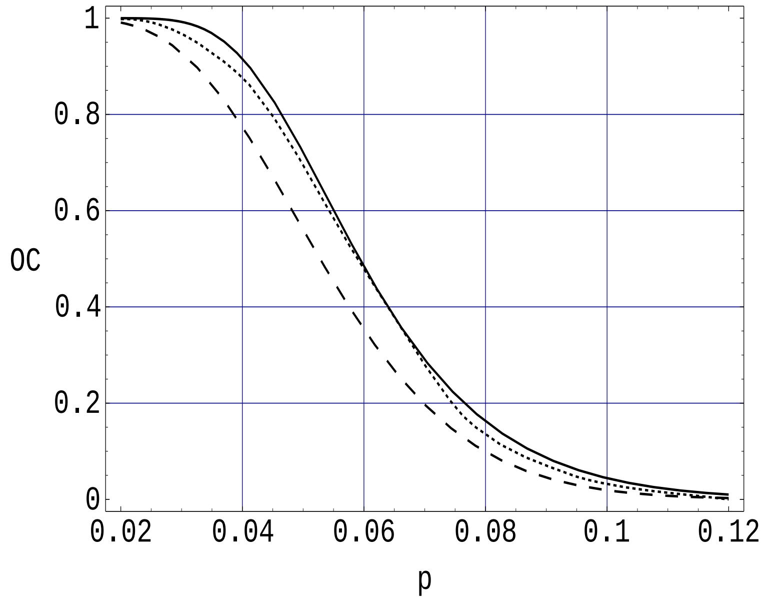
No.	p_1 / p_2	$1-\alpha$ / β	P1	P2	C	FR1	FR2	L	N	E	Δ
1	0.0521	0.95	.953	.944	.953	.955	.930*	.947	.948	.942	.991
	0.1975	0.10	.112	.066	.108	.100	.067	.107	.102	.093	.114*
2	0.0634	0.90	.922	.896	.914	.912	.880*	.909	.930	.903	.977
	0.1975	0.10	.096	.062	.084	.106	.072	.096	.110	.095	.107
3	0.0100	0.90	.935	.926	.938	.940	.932	.923	.935	.935	.979
	0.0600	0.10	.010	.086	.095	.099	.097	.098	.095	.109	.112*
4	0.0100	0.974	.976	.983	.969	.977	.976	.966	.972	.968	.995
	0.0592	0.10	.091	.093	.109	.092	.092	.098	.102	.082	.115*
5	0.0150	0.90	.934	.934	.939	.931	.925	.911	.910	.922	.964
	0.0592	0.10	.094	.070	.119*	.099	.076	.082	.102	.087	.114*
6	0.0100	0.99	.989	.984*	.989	.991	.987	.984*	.978*	.981*	.998
	0.0600	0.10	.099	.090	.127*	.109	.079	.094	.101	.106	.113
7	0.0360	0.95	.953	.935*	.939	.945	.940	.925*	.919*	.929*	.969
	0.0866	0.10	.077	.071	.104	.093	.065	.074	.094	.093	.088
8	0.0406	0.90	.908	.880*	.891	.904	.899	.877*	.879*	.869*	.931
	0.0866	0.10	.107	.063	.098	.099	.070	.084	.085	.088	.096
9	0.0100	0.99	.988	.992	.983*	.986	.988	.986	.982*	.986	.996
	0.0600	0.01	.005	.003	.007	.006	.004	.005	.004	.003	.009
10	0.0100	0.99	.991	.989	.990	.991	.991	.986	.983*	.986	.996
	0.0300	0.10	.101	.112	.115*	.104	.095	.090	.086	.091	.204*

Remarks on Table 6:

1. $1-\alpha$-values are rounded down to three decimal digits, and β-values are rounded up to three decimal digits.

2. The star (*) indicates that the nominal OC-value is not in the confidence region.

Figure 1: OC-curves for example 7



the new sampling plan was constructed for Pareto(1). This is illustrated in Table 7 for the ten examples considered. In column *simulated* $(1-\alpha)$ *worst case* also the corresponding density is presented. The entry ! indicates that the two-point condition is satisfied for all densities considered.

We assume that the sample size n is the term of interest, i.e. we are interested in keeping n as small as possible or to obtain a steep OC for a given n.

To further clarify the superiority of the new variable sampling plan over the attribute plan we consider for fixed sample sizes n, fixed p_1, p_2 and β the optimal attribute plan, i.e. the attribute plan maximizing $(1-\alpha)$.

Recall that it is assumed that the sample is drawn from a continuous population, meaning that the population is sufficiently large. Let c_A be the maximal number of nonconforming items in the sample that are accepted. The OC of the attribute plan (n, c_A) is given by

$$\begin{aligned} L_{n,c_A}(p) &= \sum_{i=0}^{c_A} \binom{n}{i} p^i (1-p)^{n-i} = 1 - \int_0^{\frac{n-c_A}{c_A+1}\frac{p}{1-p}} f_{2(c_A+1),2(n-c_A)}(y)\, dy \\ &= 1 - F_{2(c_A+1),2(n-c_A)}\left(\frac{n-c_A}{c_A+1}\frac{p}{1-p}\right), \end{aligned}$$

where, again, $F_{\nu_1,\nu_2}(y)$ is the c.d.f. of the **F**-distribution with (ν_1, ν_2) degrees

Table 7: Comparison of the new variable sampling plan with attribute plans

No.	p_1	p_2	β	nominal $1-\alpha$	n	new variable plan m	c	simulated $1-\alpha$ worst case	attribute plan c_A	$1-\alpha_A$
1	0.0521	0.1975	0.1	0.95	31	9	0.1189	0.93 (F2)	2	0.77
2	0.0634	0.1975	0.1	0.90	34	10	0.1204	0.88 (F2)	3	0.74
3	0.0100	0.0600	0.1	0.90	63	10	0.0251	!	0	0.55
4	0.0100	0.0592	0.1	0.974	82	13	0.0294	!	1	0.78
5	0.0150	0.0592	0.1	0.90	88	14	0.0396	!	2	0.83
6	0.0100	0.0600	0.1	0.99	88	14	0.0317	0.978 (N)	2	0.92
7	0.0360	0.0866	0.1	0.95	140	26	0.0593	0.919 (N)	7	0.88
8	0.0406	0.0866	0.1	0.90	145	27	0.0598	0.867 (E)	7	0.76
9	0.0100	0.0600	.01	0.99	194	31	0.0244	0.982 (N)	4	0.96
10	0.0100	0.0300	0.1	0.99	362	47	0.0204	0.983 (N)	6	0.93

of freedom and $f_{\nu_1,\nu_2}(y)$ is the corresponding density, $(0 < c_A < n)$ (cf. Uhlmann (1982), (2.38)).

The relation $L_{n,c_A}(p_2) \leq \beta$ is equivalent to

$$p_2 \geq \frac{(c_A+1)F^{-1}(1-\beta;2(c_A+1),2(n-c_A))}{n-c_A+(c_A+1)F^{-1}(1-\beta;2(c_A+1),2(n-c_A))}, \tag{25}$$

where $F^{-1}(1-\beta;\nu_1,\nu_2)$ is the quantile function of the **F**-distribution with (ν_1,ν_2) degrees of freedom (cf. Uhlmann (1982), Theorem 4.7).

The smallest number c_A satisfying the condition (25) and the corresponding $(1-\alpha_A)$ are given in the last two columns of Table 7.

The $(1-\alpha_A)$-values are considerably smaller than the worst-case simulated $(1-\alpha)$. To illustrate this property for Example 7 (cf. Table 4) the OC-curve for the attribute plan (n, c_A) is included in Figure 1 (cf. the dashed line).

For normally distributed populations, of course, the LR- or ML-sampling plans are to be preferred. But usually, there is no exact information about the distribution of the underlying population in practice.

If the underlying c.d.f. is continuous and not U-shaped, the new variable sampling plan instead of an attribute plan should be applied.

Acknowledgement

The author wishes to thank the referees, Professors W. Krumbholz and H.-J. Lenz and Dr. K. Dohmen for helpful comments.

REFERENCES

BRUHN-SUHR, M. and KRUMBHOLZ, W. (1990). A new variable sampling plan for normally distributed lots with unknown standard deviation and double specification limits. *Statistische Hefte* **31** 195-207.

BRUHN-SUHR, M. and KRUMBHOLZ, W. (1991). Exact two-sided Lieberman-Resnikoff sampling plans. *Statistische Hefte* **32** 233-241.

DUMOUCHEL, W.H. (1983). Estimating the stable index α in order to measure tail thickness: a critique. *Annals of Statistics* **11** 1019-1031.

FALK, M. (1987). Best attainable rate of joint convergence of extremes. *Extreme Value Theory* (J.Hüsler and R.-D.Reiss, eds.) 1-9. Springer, Berlin.

KÖSSLER, W., LENZ, B. and LENZ, H.-J. (1994). ExVar - Exact Variable Inspection Plans in Statistical Quality Control. *Statistical Software Newsletter, Computational Statistics and Data Analysis* **17** 97-99.

KÖSSLER, W. and LENZ, H.-J. (1995). On the robustness of Lieberman-Resnikoff sampling plans by variables. *IAPQR Transactions, Journal of the Indian Association for Productivity, Quality and Reliability* **20** 93-105.

KÖSSLER, W. and LENZ, H.-J. (1997). On the Non-Robustness of Maximum-Likelihood Sampling Plans by Variables. *Frontiers in Statistical Quality Control* **5** (H.J. Lenz and P.-Th. Wilrich, eds.) 38-51.

LIEBERMAN, G.J. and RESNIKOFF, G.J. (1955). Sampling plans for inspection by variables. *Journal of the American Statistical Association* **50** 457-516.

MASUDA, K. (1978). Effect of Non-Normality on sampling plans by Lieberman and Resnikoff. *International Congress of Quality Control, Tokyo* D3-7 - D3-11.

PICKANDS, J. (1978). Statistical inference using extreme order statistics. *Annals of Statistics* **3** 119-131.

PP NUMATH2 (1984). *Programmpaket für Verfahren der numerischen Mathematik, User's Guide*, Humboldt-Universität zu Berlin.

REISS, R.-D. (1989). *Approximate Distributions of Order Statistics*, Springer, New York.

RESNIKOFF, G.J. (1952). A new two-sided acceptance region for sampling by variables. *Technical Report* **8**, Applied Mathematics and Statistics

Laboratory, Stanford University.

SCHNEIDER, H. and WILRICH, P.-TH. (1981). The Robustness of Sampling Plans for Inspection by Variables. *Computational Statistics, Festschrift dedicated to W.Wetzel* 281-295.

SMITH, R.L. (1987). Estimating tails of probability distributions. *Annals of Statistics* **15** 1174-1207.

UHLMANN, W. (1982). *Statistische Qualitätskontrolle*, Teubner, Stuttgart.

Wolfgang Kössler
Institut für Informatik
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin