Variable inspection plans for continuous populations with unknown short tail distributions

Wolfgang Kössler

Abstract The ordinary variable inspection plans are sensitive to deviations from the normality assumption. A new variable inspection plan is constructed that can be used for arbitrary continuous populations with short tail distributions. The peaks over threshold method is used, the tails are approximated by a generalized Pareto distribution, their parameters and the fraction defective are estimated by a moment method proposed in a similar form by Smith and Weissman (1985). The estimates of the fraction defective are asymptotically normal. It turns out that their asymptotic variances do not differ very much for the various distributions. Therefore we may fix the variance and use the known asymptotic distribution for the construction of the inspection plans. The sample sizes needed to satisfy the two-point conditions are much less than that for attribute plans.

1 Introduction

We consider a lot of units having a quality characteristic $\mathbf{X}$ with a (unknown) continuous cumulative distribution function (cdf) F. Given a sample $X_1,...,X_n$ a decision is to be made whether the lot is to be accepted or not. For simplicity we assume only lower specification limits L, but the procedure can be extended to the two-sided case of lower and upper specification limits. The fraction defective p_L of the lot is defined by

$$p_L = P(\mathbf{X} < L) = F(L).$$

We intend to construct reasonable estimates $\hat{p}$ of p based on the sample. Our variable inspection plan is then defined by: If $\hat{p} \leq c$ the lot will be accepted else it will be rejected. Denote by

Wolfgang Kössler
Humboldt-Universität zu Berlin, Germany, Institute of Computer Science, e-mail: koessler@informatik.hu-berlin.de

$$L^{n,c}(p) := P_p(\hat{p} \leq c), \quad 0 < p < 0.5$$

the operating characteristic (OC). Variable inspection plans (n,c) are computed by minimizing the sample size n while meeting the so called two-point conditions ($0 < p_1 < p_2 < 1$, $0 < \beta < 1-\alpha$)

$$L^{n,c}(p_1) \geq 1-\alpha \qquad \text{and} \qquad L^{n,c}(p_2) \leq \beta, \tag{1}$$

where p_1 and p_2 are the accepted and rejected quality level, respectively. The ordinary variable inspection plan (ML plan, cf. e.g. Uhlmann, 1992, for the two-sided case see Bruhn-Suhr and Krumbholz, 1990) is very sensitive with respect to deviations from the normal distribution assumption (cf. Kössler and Lenz, 1995,1997).

In this paper we construct variable inspection plans which do not rely on the normality assumption and which require less sample sizes than the attribute plan. The main idea is that nonconforming items X_i occur in the lower tail of the underlying cdf, namely $X_i < L = F^{-1}(p_L)$ with the (unknown) fraction defective $p = p_L$. Additionally, items X_i with $X_i \approx L, X_i > L$ can be considered suspicious. They also should be considered in inspection plans. Whereas in Kössler (1999) we assumed that the underlying density has not too short tails to obtain Maximum Likelihood estimates we consider the short tail case here and use a (moment) estimate proposed by Smith (1987) and Smith and Weissman (1985).

In section 2 we apply the peak over threshold method, approximate the tails of the density by a generalized Pareto distribution (GPD) and estimate their parameters and the fraction defective in section 3. Using the asymptotic normality of all these estimators we compute inspection plans meeting the conditions (1) at least approximately in section 4. Comparisons of the various sampling plans in section 5 show that the necessary sample sizes for the new plan are much less than that for the attribute sampling plan. Simulation studies performed in section 6 show that this method works quite well even for relatively small sample sizes.

2 Approximation of the tails by a GPD

We assume that we have a short tail density with a lower endpoint x_0 which may be set to zero without restriction to the generality, more precisely, we assume that the underlying density is in the domain of attraction of the Weibull (cdf. $G_\gamma(x) = 1 - exp(-x^\gamma)$, tail index $\gamma > 0$, $x > 0$). Let $t = t_L$ be a lower threshold value and $0 < y < t$. The conditional cdf $F_t(y)$ of $t - X$ conditioned under $X < t$,

$$F_t(y) = \frac{F(t) - F(t-y)}{F(t)}, \tag{2}$$

can be approximated by a generalized Pareto cdf,

$$GPD(y;\sigma,k) := 1-\left(1-\frac{ky}{\sigma}\right)^{\frac{1}{k}} \qquad k,\sigma>0, \quad 0<y<\frac{\sigma}{k}$$

as was shown by Pickands (1975, Theorem 7). The parameters $k=\frac{1}{\gamma}>0$ and $\sigma(t)=k(t-x_0)=kt$ are given by the extreme value distribution theory, cf. eg. Falk (1987).

To approximate the fraction defective p_L let $t=t_L$ be fixed, $t>L=F^{-1}(p_L)$ and $y=y_L=t-L$. We obtain from (2):

$$p_L = F(L) = F(t) - F_t(y)\cdot F(t) \approx F(t)\left(1-\frac{ky}{\sigma}\right)^{\frac{1}{k}}.$$

3 Estimation of the fraction defective

Define the threshold by $t_L=F^{-1}(q)$ for given q, $0<q<0.5$ and estimate it by $\hat{t}_L = X_{(m+1)}$, where $m:=\lfloor nq\rfloor$ and $X_{(i)}$ is the ith order statistics of the sample. Let $\hat{y}_L=\hat{t}_L-L$ and $(\hat{k}_L,\hat{\sigma}_L)$ be a consistent estimate of (k_L,σ_L) in the GPD-model. Then

$$\hat{p}_L = q\cdot\begin{cases}\left(1-\frac{\hat{k}_L\hat{y}_L}{\hat{\sigma}_L}\right)^{\frac{1}{\hat{k}_L}} & \text{if} \quad \hat{k}_L\neq 0\\ e^{-\frac{\hat{y}_L}{\hat{\sigma}_L}} & \text{if} \quad \hat{k}_L=0\end{cases} \tag{3}$$

is a consistent estimate of p_L.

Note that $\hat{y}_L$ is random, and the estimate (3) is well defined if $\hat{y}_L\geq 0$ and if

$$\hat{k}_L\hat{y}_L < \hat{\sigma}_L. \tag{4}$$

In the few cases that $\hat{y}_L<0$ we may reject the lot without further computations because these cases indicate low quality.

For the estimation of the parameters we might use Maximum Likelihood estimates. This procedure was pursued in Kössler (1999). However, if $k>0.5$ the ML estimates are not asymptotically normal ($k\leq 1$) or they do not exist ($k>1$). For the short tail densities here we use an estimate (SW estimate) proposed by Smith and Weissman (1985), eq. (4.3) and Smith (1987), section 7,

$$\hat{k}_L = \frac{1}{m}\sum_{i=2}^{m}\log\frac{X_{(m+1)}-X_{(1)}}{X_{(i)}-X_{(1)}} \tag{5}$$

$$\hat{\sigma}_L = \hat{k}_L(X_{(m+1)}-X_{(1)}). \tag{6}$$

The estimate for k_L may be motivated by the moment equation $E\left(-\log(1-\frac{kY}{\sigma})\right)=k$ if Y is a random variable, $Y\sim GPD(\sigma,k)$, cf. Smith (1987). The estimate for σ_L is motivated by $\sigma_L=k_L(t_L-x_0)$ if $k_L>0$, and the in praxis unknown lower endpoint x_0 is estimated by the smallest observation $X_{(1)}$.

If $X_{(1)}<L$ then condition (4) is satisfied and consistent estimates $\hat{p}_L$ of the fraction defective p_L are obtained by inserting $\hat{k}_L$ and $\hat{\sigma}_L$ in (3). In the case of $X_{(1)}\geq L$

we may set $\hat{p}_L := 0$ as it indicates good quality. However, a slight negative bias may be introduced.

Note that we also investigated various other estimates of (k,σ), moment estimates (MOM), probability weighted moment estimates (PWM) and elemental percentile moment estimates (EPM) (see e.g. Beirlant (2004) and references therein). The asymptotic variances of the MOM and PWM estimates are much larger than that of estimates (5) and (6), a result that is confirmed by finite sample simulation studies. Simulations with the EPM method show that their bias is slightly less than that for the SW estimate but the variances are much higher for the EPM estimates.

Under certain conditions on the convergence of $t \to x_o$, $L \to x_o$ if $n \to \infty$ the SW-estimate $\hat{p}_L$ is asymptotically normally distributed with expectation zero and variance $V(p_L)$,

$$\sqrt{m}\frac{\hat{p}_L - p_L}{p_L} \to \mathcal{N}\big(0, V(p_L)\big) \tag{7}$$

(cf. Smith, 1987, ch.8). To obtain a closed relation for the variance $V(p_L)$ dependent on the cdf F we follow the arguments of Smith (1987, ch.8).

Let $z,\ z > 0$, be fixed and define the sequences $p_m, q_m, \quad p_m \to 0,\ q_m \to 0, \quad 0 < p_m < q_m$ in the same way as in Smith (1987) by

$$z = 1 - \frac{k y_m}{\sigma_m} = 1 - \frac{k(t_{L,m} - L_m)}{\sigma_m} = 1 - \frac{k(F^{-1}(q_m) - F^{-1}(p_m))}{\sigma_m}, \tag{8}$$

where $y_m := t_{L,m} - L_m = F^{-1}(q_m) - F^{-1}(p_m)$, and k and σ_m are the parameters given by the GPD approximation of the conditional probability (2) which depend on the sequence of the threshold values $t_{L,m} = F^{-1}(q_m)$.

The asymptotic variance $V = V_F$ for $p_m, q_m \to 0$ is then given by

$$V_F = 1 - q_m + \mathbf{c}^T\mathbf{S}\mathbf{c}, \tag{9}$$

where

$$\mathbf{c}^T = (-\frac{1}{k}(\frac{1}{z} - 1), \frac{\log z}{k^2} + \frac{1}{k^2}(\frac{1}{z} - 1)) \qquad \text{and} \qquad \mathbf{S} = \begin{pmatrix} 1 & k \\ k & k^2 \end{pmatrix}. \tag{10}$$

The term $\mathbf{c}^T\mathbf{S}\mathbf{c}$ in (9) becomes $\mathbf{c}^T\mathbf{S}\mathbf{c} = \frac{1}{k^2}\log^2 z$, where z is defined by (8) with $\sigma_m = k t_{L,m}$. Interestingly, if the cdf F is GPD or Weibull the term $\mathbf{c}^T\mathbf{S}\mathbf{c}$ is independent of k.

For further investigation of the variance term we considered the following short tail densities, the GPD, the Weibull, the Beta, the Gamma and the Burr, all with various values of the parameter k. From the matrices $\mathbf{S}$ in the ML and SW cases (cf. Smith, 1987) it may be seen that for $k = 0.5$ the asymptotic variances of the ML and SW estimates are the same. Moreover, it turns out that the asymptotic variances of the SW estimates of $\hat{p}_L$ in the case of short tails are often similar to that of the ML estimate in the (long tail) Pareto ($k = -1$) case where we had an upper specification limit (Kössler, 1999, eq. (14), Table 3). Exceptions are the Gamma with $k = 0.25$, and the Burr with $k = 0.25$ or $k = 0.5$. An explanation of the latter facts may be that the speed of convergence of $q_m \to 0$ must be faster in that cases, cf. convergence conditions SR1 or SR2 of Smith (1987).

However, if the ratios of the fraction defective p_L and the used tail fraction q are not too small then the dissimilarities are not so large.

The similarities of the asymptotic variances will allow us to use the variance $V(p)$ obtained for the ML estimate in the (reverse) Pareto ($k = -1$) case for the determination of the sampling plan later on. This variance is given by (9) but with $\mathbf{S} = (1-k)\begin{pmatrix} 2 & 1 \\ 1 & 1-k \end{pmatrix} = 2\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ (cf. eg. Smith, 1987).

4 The new sampling plan

Since we have established asymptotic normality with similar variances for the various underlying cdfs we may proceed in the same way as in Kössler (1999) to determine a new sampling plan. Given $q > 0$ define $m = \lfloor nq \rfloor$, i.e. for the estimation of the fraction defective only the $m+1$ smallest observations are used.

For a discussion of the choice of the threshold values t_L we refer to Kössler (1999). Here we apply a slightly modified version

$$q = q(n_0) = p_2 + \frac{1}{\sqrt{n_0}},$$

where n_0 is an initial estimate of the sample size, $n_0 = \frac{n_V + n_A}{2}$, n_V and n_A are the sample sizes for the ordinary variable sampling and for the attribute sampling plan, respectively. This definition reflects the conditions $q \to 0$, $q > p$ and also the fact that the resulting sample size is expected to lie between n_V and n_A.

Since the number 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 $\hat{p}$.

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

$$L^{n,c}(p_1) \approx \Phi\big(\sqrt{m}\frac{c - p_1}{p_1\sqrt{V(p_1)}}\big) = 1 - \alpha, \qquad L^{n,c}(p_2) \approx \Phi\big(\sqrt{m}\frac{c - p_2}{p_2\sqrt{V(p_2)}}\big) = \beta.$$

An to integer values for m adjusted solution (m, c) of this system of equations is given by

$$m = \left\lceil \frac{1}{(p_1 - p_2)^2}\big(p_2\sqrt{V(p_2)}\Phi^{-1}(\beta) - p_1\sqrt{V(p_1)}\Phi^{-1}(1-\alpha)\big)^2 \right\rceil \qquad (11)$$

$$c = p_1 + \Phi^{-1}(1-\alpha)\frac{p_1\sqrt{V(p_1)}}{\sqrt{m}}. \qquad (12)$$

Given the numbers m and q the sample size n is determined by $m = \lfloor nq \rfloor$. In such a way a new sampling plan (n, m, c) is obtained. It is given by

$$n = \left\lceil \frac{m}{q} \right\rceil, \qquad \text{where} \qquad q = p_2 + \frac{1}{\sqrt{n_0}}$$

and m and c are given by (11) and (12).

Note that $V(p)$ is also dependent on q, and since the definition of q is slightly modified, these variances and also the sample sizes are slightly different from that in Kössler (1999).

First simulations show that the OC estimates are slightly shifted to the right. Therefore, the acceptance number is empirically modified to $c_{SW} := c \cdot (1 - 1/n)$.

5 Comparison with other sampling plans

In Table 1 the sampling plans (n, m, c_{SW}) for twelve different two-point conditions are presented. For comparison the corresponding sample sizes n_V and n_A of the ML-variable sampling plan and the attribute sampling plan, respectively, are given in the last two columns of Table 1. The sample sizes for the ordinary ML variable sampling plan are computed by the R program ExLiebeRes of Krumbholz and Steuer (2014). Since our new sampling plan can be used also in the case of two-sided specification limits the sample size n_V is computed for that case.

Table 1 The new sampling plan (n, m, c_{SW}) together with the sample sizes n_V and n_A of the ordinary variable sampling plan and the attribute sampling plan, respectively.

	two-point condition				new sampling plan				
No.	p_1	$1-\alpha$	p_2	β	n	m	c_{SW}	n_V	n_A
1	0.0521	0.9500	0.1975	0.10	31	11	0.1053	27	45
2	0.0634	0.9000	0.1975	0.10	31	11	0.1072	27	45
3	0.0100	0.9000	0.0600	0.10	59	11	0.0237	36	88
4	0.0100	0.9743	0.0592	0.10	80	13	0.0280	54	133
5	0.0152	0.9000	0.0592	0.10	83	14	0.0292	54	111
6	0.0100	0.9900	0.0600	0.10	90	14	0.0303	64	153
7	0.0360	0.9500	0.0866	0.10	143	24	0.0576	106	189
8	0.0406	0.9000	0.0866	0.10	149	25	0.0581	107	189
9	0.0100	0.9900	0.0600	0.01	203	27	0.0237	111	263
10	0.0200	0.9500	0.0500	0.05	316	34	0.0309	186	410
11	0.0100	0.9900	0.0300	0.10	390	33	0.0198	217	590
12	0.0200	0.9900	0.0300	0.01	4609	213	0.0244	2241	5362

The examples 1,2,4,5,7 and 8 are from Resnikoff (1952), example 10 is from Steland and Zähle (2009). From Table 1 it can be seen that the sample sizes for the new plan are considerably less than that for the attribute sampling plan.

6 Simulation study

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 (reverse) Pareto with $k=-1$ (where ML estimates are used).

To investigate whether the sampling plans constructed can be applied for short tail densities as well as for moderate sample sizes simulation studies are carried out. The OC is estimated for the same examples as in the previous section. Note that the examples 1-9, 11 are the same as in Kössler (1999) but the necessary sample sizes may differ slightly since the definition of the used fraction of the sample is altered. To see whether the asymptotic theory works in practice we have included an example with very large sample sizes (Example 12).

The simulation size is $M=2000$. The following c.d.fs are included in the simulation study: GPD, Weibull, Gamma, and Burr, all with with $k=0.25, 0.5, 0.75, 1$.

We obtain that for $k \geq 0.5$ and for most densities the OC is well estimated. To give an impression on the goodness of the estimated OC values we present only the worst cases. For β they are $\hat{\beta}=0.13$ (instead of $\beta=0.10$). For $1-\alpha$ they are $1-\hat{\alpha}=0.97$ (instead of $1-\alpha=0.99$),$1-\hat{\alpha}=0.93$ (instead of $1-\alpha=0.95$), $1-\hat{\alpha}=0.86$ (instead of $1-\alpha=0.90$). For $k=0.25$ the estimates are only slightly worse. Perhaps somewhat surprisingly, the latter happens also in example 12 where we have very large sample sizes. Note that, for $k=0.25$ a ML estimate is to be preferred.

7 Adaptive procedure and summary

In the short tail case the estimates of the fraction defective are different from that in the medium or long tail case. Since it is generally not known which case occurs, we suggest to apply an adaptive procedure. First the sample size is determined in the way described. Then, after the sample is drawn from the lot, the parameter k_L is estimated by the SW method. If $\hat{k}_L=\hat{k}_{L,SW} \leq 0.5$ we assume that we have a medium or long tail, estimate k_L, σ_L and the fraction defective p_L by the ML method and use the modified acceptance number c_{ML} from the ML plan (cf. Kössler,1999). If $\hat{k}_{L,SW}>0.5$ we assume that we have a short tail, estimate k_L, σ_L and the fraction defective p_L by the SW method and use the modified acceptance number c_{SW}.

For normally distributed populations, of course, the ML-sampling plans are to be preferred. But usually, there is no exact information about the distribution of the underlying population in practice. Therefore, if the underlying c.d.f. is continuous the new variable sampling plan instead of an attribute plan should be applied. If it is known that we have short tails the sampling plan proposed here should be applied.

References

1. Beirlant J, Goegebeur, Y, Segers J and Teugels J (2004) Statistics of Extremes, Theory and Applications. Wiley & Sons, Chichester
2. 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
3. Falk M (1989) Best attainable rate of joint convergence of extremes. In: Hüsler J, Reiss R-D (eds) Extreme value theory, Proceedings of a conference held in Oberwolfach, Dec.6-12, 1987, pp. 1-9, Springer
4. Kössler W (1999) A new One-sided variable inspection plan for continuous distribution functions. Allgemeines Statistisches Archiv 83: 416-433
5. Kössler W and Lenz H-J (1995) On the Robustness of Lieberman-Resnikoff Sampling Plans by Variables. Journal of the Indian Association for Productivity, Quality and Reliability 20:93-105
6. Kössler W and Lenz H-J (1997) On the Non-Robustness of Maximum-Likelihood Sampling Plans by Variables In: Lenz, H-J and Wilrich P-Th (eds) Frontiers in Statistical Quality Control 5, pp. 38-52, Physica, Heidelberg
7. Krumbholz W and Steuer D (2014) On exact and optimal single sampling plans by variables. Advances in Statistical Analysis 98:87-101
8. Pickands J (1975) Statistical inference using extreme order statistics. Annals of Statistics 3:119-135
9. Resnikoff G V (1952) A new two-sided acceptance region for sampling by variables. Technical Report 8 Applied mathematics and Statistics Laboratory, Stanford University
10. Smith R L (1987) Estimating tails of probability distributions. Annals of Statistics 15:1174-1207
11. Smith R L and Weissman I (1985) Maximum Likelihood estimation of the lower tail of a probability distribution. Journal of the Royal Statistical Society B 47:285-298
12. Steland A and Zähle H (2009) Sampling inspection by variables: nonparametric setting. Statistica Neerlandica 63:101-123
13. Uhlmann W (1982) Statistische Qualitätskontrolle. Teubner, Stuttgart