UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Univariate parametric and nonparametric statistical quality control techniques with estimated process parameters

by

Schalk William Human

Submitted in partial fulfillment of
the requirements for the degree

Philosphiae Doctor (Mathematical Statistics)

in the
Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria

July 2009

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Declaration

I declare that the thesis that I hereby submit for the degree Philosphiae Doctor (Mathematical Statistics) at the University of Pretoria has not previously been submitted by me for degree purposes at any other university.

Signature ____________________ Date ________________

Acknowledgements

Thanks go to many for making this research and degree possible.

I would like to sincerely thank my supervisor Prof. CF Smit and my co-supervisor Prof. S Chakraborti for their extraordinary efforts to my benefit throughout my tenure as a PhD student in the Department of Statistics at the University of Pretoria. Their professional knowledge, constructive guidance and years long experience were indispensable during the entire process of doing my research and preparing this dissertation.

I am truly grateful to my colleagues at the Department of Statistics at the University of Pretoria. I have thoroughly enjoyed collaborating and learning from you. Thank you for every bit of your support, encouragement and kindness.

I am also indebted to Prof. NAS Crowther for allowing me in the PhD program and his unconditional trust for granting me permission to take research leave, and STATOMET along with the NRF for supporting me financially. Thank you very much.

Thanks also go to the faculty members and the postgraduate students of the Department of Information Systems, Statistics and Management Science at the University of Alabama, USA. They made my research visits a real pleasure.

Special thanks to my family – especially my parents Will and Elsabé, and my brother Cornel. Your lifelong love, sacrifices and continuous support has meant the world to me!

My Creator and Saviour for the talents and grace bestowed upon me.

Summary

Chapter 1 gives a brief introduction to statistical quality control (SQC) and provides background information regarding the research conducted in this thesis.

We begin **Chapter 2** with the design of Shewhart-type Phase I S^2, S and R control charts for the situation when the mean and the variance are both unknown and are estimated on the basis of m independent rational subgroups each of size n available from a normally distributed process. The derivations recognize that in Phase I (with unknown parameters) the signaling events are dependent and that more than one comparison is made against the same estimated limits simultaneously; this leads to working with the joint distribution of a set of dependent random variables. Using intensive computer simulations, tables are provided with the charting constants for each chart for a given false alarm probability. Second an overview of the literature on Phase I parametric control charts for univariate variables data is given assuming that the form of the underlying continuous distribution is known. The overview presents the current state of the art and what challenges still remain. It is pointed out that, because the Phase I signaling events are dependent and multiple signaling events are to be dealt with simultaneously (in making an in-control or not-in-control decision), the joint distribution of the charting statistics needs to be used and the recommendation is to control the probability of at least one false alarm while setting up the charts.

In **Chapter 3** we derive and evaluate expressions for the run-length distributions of the Phase II Shewhart-type *p*-chart and the Phase II Shewhart-type *c*-chart when the parameters are estimated. We then examine the effect of estimating p and c on the performance of the *p*-chart and the *c*-chart via their run-length distributions and associated characteristics such as the average run-length, the false alarm rate and the probability of a "no-signal". An exact approach based on the binomial and the Poisson distributions is used to derive expressions for the Phase II run-length distributions and the related Phase II characteristics using expectation by conditioning (see e.g. Chakraborti, (2000)). We first obtain the characteristics of the run-length distributions conditioned on point estimates from Phase I and then find the unconditional characteristics by averaging over the distributions of the point estimators. The in-control and the out-of-

control properties of the charts are looked at. The results are used to discuss the appropriateness of the widely followed empirical rules for choosing the size of the Phase I sample used to estimate the unknown parameters; this includes the number of reference samples m and the sample size n.

Chapter 4 focuses on distribution-free control charts and considers a new class of nonparametric charts with runs-type signaling rules (i.e. runs of the charting statistics above and below the control limits) for both the scenarios where the percentile of interest of the distribution is known and unknown. In the former situation (or Case K) the charts are based on the sign test statistic and enhance the sign chart proposed by Amin et al. (1995); in the latter scenario (or Case U) the charts are based on the two-sample median test statistic and improve the precedence charts by Chakraborti et al. (2004). A Markov chain approach (see e.g. Fu and Lou, (2003)) is used to derive the run-length distributions, the average run-lengths, the standard deviation of the run-lengths etc. for our runs rule enhanced charts. In some cases, we also draw on the results of the geometric distribution of order k (see e.g. Chapter 2 of Balakrishnan and Koutras, (2002)) to obtain closed form and explicit expressions for the run-length distributions and/or their associated performance characteristics. Tables are provided for implementation of the charts and examples are given to illustrate the application and usefulness of the charts. The in-control and the out-of-control performance of the charts are studied and compared to the existing nonparametric charts using criteria such as the average run-length, the standard deviation of the run-length, the false alarm rate and some percentiles of the run-length, including the median run-length. It is shown that the proposed "runs rules enhanced" sign charts offer more practically desirable in-control average run-lengths and false alarm rates and perform better for some distributions.

Chapter 5 wraps up this thesis with a summary of the research carried out and offers concluding remarks concerning unanswered questions and/or future research opportunities.

Contents

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA