SEMIPARAMETRIC INFERENCES IN STABILITY DESIGN

By

YOUNGKYOUNG MIN

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN STATISTICS

UNIVERSITY OF FLORIDA

2004

Copyright 2004

by

Youngkyoung Min

ACKNOWLEDGMENTS

I am thankful to the faculty of the Statistics Department. Each one of them has helped me throughout my education in his/her own ways and also I am thankful to the faculty for giving me the opportunity to study statistics at the University of Florida and for supporting me to finish my study. I would especially like to thank Dr. Ying Qing Chen for his support, time, and valuable guidance. I would like to thank Dr. Annpey Pong for her advice. I am greatly thankful to Dr. Cynthia W. Garvan, chair of my committee and the rest of my thesis committee, Dr. Rongling Wu and Dr. Lili Tian, for their precious time and contribution.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science in Statistics

SEMIPARAMETRIC INFERENCES IN STABILITY DESIGN

By

Youngkyoung Min

August 2004

Chair: Cynthia W. Garvan
Major Department: Statistics

Semiparametric regression models have been proposed and used in stability analysis to estimate the shelf life of drug product.

In this thesis, a new estimation approach based on the comparability of random variation of continuous repeated outcomes is proposed. The new approach will be studied under the Food and Drug Administration's (FDA) full sampling plan, and the reduced sampling plans such as matrixing and bracketing designs.

The methodologies are demonstrated by simulation studies and applications in practical data analysis.

CHAPTER 1
INTRODUCTION

In the United States, the Food and Drug Administration (FDA) requires the shelf life for every drug product in the market. The shelf life of a drug product is defined as the time interval in which the characteristic of the drug product remains within the approved United States Pharmacopedia and National Formulary (USP/NF) specification after being manufactured. The shelf life is also referred to as an expiration-dating period. A drug's shelf life is established through a stability study under the appropriate storage conditions. The purpose of a stability study is not only to characterize the degradation of a drug substance[1] or drug product,[2] but also to establish a shelf life applicable to all future batches manufactured and packaged under similar circumstances, based on testing a minimum of three batches of the drug product.

Stability study includes testing which provides evidence on how the quality of a drug substance or drug product varies with time under the influence of a variety of environmental factors, such as temperature, humidity, and light. The choice of test conditions is based on an analysis of the effects of climatic conditions in the three regions of the EU, Japan, and the United States. Information on the stability of the drug substance is an integral part of the systematic approach to stability evaluation, according to the

[1] The unformulated drug substance that may subsequently be formulated with excipients (anything other than the drug substance in the dosage form) to produce the dosage form

[2] The dosage form in the final immediate packaging intended for marketing

FDA's guidance documents. The procedure of a stability testing for the drug substance is similar with that for the drug product. The procedure of the drug substance includes Stress Testing, Selection of Batches, Container Closure System, Specification, Testing Frequency, Storage Conditions, Stability Commitment, and Evaluation. The design of the formal stability studies for the drug product should be based on knowledge of the behavior and properties of the drug substance, results from stability studies on the drug substance, and experience gained from clinical formulation studies. In case of the procedure for the drug product, Photostability Testing and Statements/Labeling are added to that for the drug substance.

A systematic approach should be adopted in the presentation and evaluation of the stability information. An appropriate statistical method should be employed to analyze the long-term primary stability data in an original application and to test the goodness of fit on all batches and combined batches to the assumed degradation line or curve. The purpose of this analysis is to establish shelf life during which a quantitative attribute will remain within acceptance criteria for all future batches manufactured, packaged, and stored under similar circumstances with a high degree of confidence. This same method could also be applied to commitment batches to verify or extend the originally approved retest shelf life.

An appropriate approach to shelf life estimation is to analyze a quantitative attribute by determining the earliest time at which the 95% confidence limit for the mean around the regression curve intersects the proposed acceptance criterion. The shelf life is usually estimated as the time at which the 95% one-sided lower confidence bound for the mean degradation curve intersects the approved specification limit if the drug

characteristic is expected to decrease with time in a linear form to determine the shelf life of a given batch of a drug product as illustrated in Figure 1.

The FDA and International Conference on Harmonization (ICH) stability guidelines require that at least three batches and preferably more batches be tested for the establishment of single shelf life for the drug product. It is suggested that stability testing be performed at three-month interval for the first year, a six-month interval for the second year and yearly after that. Stability test results can be combined for the establishment of a single shelf life if there is no batch-to-batch variation. A preliminary test for batch similarity should be performed at the 25% level of significance before the data can be combined for analysis.

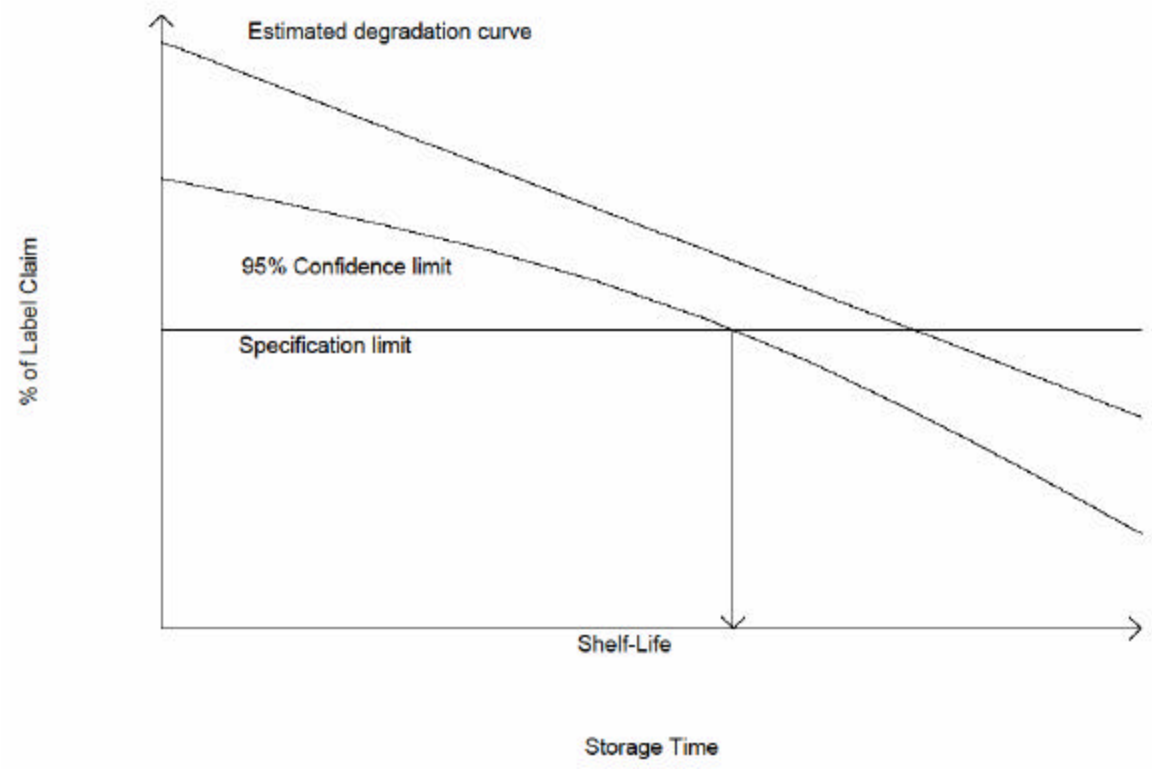

Figure 1. Illustration of Shelf Life

Regression Analysis is considered an approach to evaluate the stability data and establish a shelf life. The nature of the relationship between an attribute and time will

determine whether data should be transformed for linear regression analysis. Usually, the relationship can be represented by a linear or non-linear function on an arithmetic or logarithmic scale. Sometimes a non-linear regression can be expected to better reflect the true relationship.

In both research and industrial field, the linear regression models have been used extensively. When the batch to batch variation of the product tested is not concerned, the following simple linear regression model can be used to describe the continuous characteristic over time: $Y = \alpha + \beta X + e$, where α and β are unknown parameters, X is the fixed measurement time determined by study design, and e is the zero mean normal deviate. When the batch-to-batch variation is concerned, the random effects models are used. α and β varies among batches according to some parametric bivariate normal distribution (Chow and Shao, 2001).

Most of the linear regression models in stability study have been used with the parametric assumptions. In fact, it may not be always true that these parametric assumptions are appropriate. For example, when the number of manufacturing batches is actually limited for the drug product, then the continuous bivariate normal assumption on α and β is doubt in dealing with the batch-to-batch variations. In 2003, Chen, Pong and Xing used a semiparametric rank regression approach with a linear model with unspecified error distribution, not the assumption of the error normality. They showed that the semiparametric regression models and their associated rank-based inference procedures in obtaining the self life provide an alternative approach when the parametric assumptions are not necessarily true in the fully parametric models and the feature of the non parametric components also have profound practical interpretation.

A full study design is one in which samples for every combination of all design factors are tested at all time points. A reduced design is one in which samples for every factor combination are not all tested at all time points. A reduced design can be a suitable alternative to a full design when multiple design factors are involved. Any reduced design should have the ability to adequately predict the retest period or shelf life. The ICH and FDA stability guidelines recommend that a reduced stability design such as the bracketing design and the matrixing design be considered as an alternative to the full factorial design to avoid costly and time consuming. A bracketing design is one where the samples on the extremes of ordered levels of an appropriate factor are tested. A matrixing design is one where a fraction of the total number of samples is tested at any specified sampling point (FDA, ICH, 2002). The underlying assumption of both bracketing and matrixing designs is that the stability of the samples tested is representative to those that are not tested.

In 2001, Annpey Pong compared the performances of bracketing and matrixing designs in stability studies, in terms of their powers for detection of a scientifically meaningful difference between the rate of stability loss and the precision of the estimated drug shelf life. It was shown that the bracketing design is a better design than the matrixing design.

In this thesis, we proposed and studied the semiparametric regression models in stability analysis to estimate the shelf life of drug product and a new estimation approach based on the comparability of random variation of continuous repeated outcomes. The methodologies are illustrated by simulation studies and applications in practical data analysis.

The remaining part of this thesis is organized as follows. In the next chapter, a literature review includes regulatory requirement, stability design and statistical methods. The semiparametric model, pair-wise estimation and determination of shelf life are presented in Chapter 3. In Chapter 4, numerical studies including simulation and data examples are illustrated. In Chapter 5, stability analysis on the difference between full and reduced (bracketing; matrixing) designs by pair-wise estimation is discussed. The direction for future research in this field is given in Chapter 6.

CHAPTER 2
REVIEW OF LITERATURE

In this chapter, the background, regulatory requirements, several kinds of stability design, and statistical methods are illustrated.

Regulatory Requirements

In 1962, a significant Kefauver-Harris Drug Amendment was passed, which not only strengthened the safety requirements for new drugs, but also established an efficacy requirement for new drugs for the first time. In 1984,Congress passed the Price Competition and Patent Term Restoration Act to provide an increased patent protection to compensate for patent life last during the approval process. Based on this act, the FDA was authorized to approve generic drug only based on bioavailability and bioequivalence trials on healthy subjects. In 1993, the ICH issues guidelines on stability based on a strong industrial interest in international harmonization of regulatory requirements for marketing in European Communities (EC), Japan and the USA. A draft FDA stability guideline was revised and expanded in 1998. The purpose of this section is to provide an overview of the regulatory requirement for stability testing. The requirements from the FDA and the ICH guidelines are following below:

FDA Guideline

The FDA issued a guideline for the stability of human drugs and biologics under part 21 Code of Federal Regulations (CFR) Section 10.9, in 1998. The purpose of this stability guideline is to provide not only recommendations for the design and analysis of stability studies for establishing an appropriate expiration-dating period, but also

recommendations for submission of stability information for investigational new drug (IND) application, new drug application (NDA), abbreviated new drug application (ANDA) and biological product license application (PLA). The FDA stability guideline suggests that the expiration-dating period be estimated as the time at which the 95% one-sided lower confidence bound for the mean degradation curve intersects the approved lower specification limit if the drug characteristic is expected to decrease over time, to determine the expiration-dating period of a given batch of a drug product (FDA, 1997).

ICH Guideline

The ICH guideline provides a general indication on the requirements for stability testing, but leaves sufficient flexibility to encompass the variety of different situations and characteristics of the materials being evaluated (Chow and Pong, 1995). For the evaluation of stability data, the ICH guideline indicates that statistical methods should be employed to test the goodness of fit of the data on all batches and combined batches to the assumed degradation line or curve. If it is inappropriate to combine data from several batches, the overall re-test period may depend on the minimum time a batch may be expected to remain within acceptable limits (FDA, ICH, 2002).

Stability Design

The FDA stability guideline provides some design considerations for stability studies under ambient conditions. These design considerations include batch sampling, container-closure and drug product sampling, and sampling time considerations. An appropriate stability design can help to achieve the objective of a stability study. Basically, a stability design consists of two parts: the selection of design factors (batch, strength and package type) and the choice of sampling intervals (three-month during the

first year, six-month during the second year, and yearly thereafter). For the selection of design factors, the stability design commonly employed is the full factorial design.

A typical stability design is a full factorial design. However, in many cases, most pharmaceutical companies are unable to conduct a full factorial design with sampling time intervals suggested by FDA due to limited resources. A reduced stability design such as a bracketing or a matrixing design with less sampling time points is usually considered as an alternative to achieve certain degree of precision without losing much information. The bracketing design and matrixing design are outlined below:

Bracketing Design

As defined in the glossary to the parent guidance (FDA, ICH Q1D, 2002), bracketing is the design of stability schedule such that only samples on the extremes of certain design factors (e.g., strength, container size and/or fill) are tested at all time points as in a full design. The design assumes that the stability of any intermediate levels is represented by the stability of the extremes tested. That is to say, the degradation rates for different levels of a factor in stability studies lie between the degradation rates of the extreme levels. The bracketing design assumes that the stability of the intermediate condition samples is included in those at the extremes.

Design factors are variables (e.g., strength, container size and/or fill) to be evaluated in a study design for their effect on product stability. Bracketing can be applied to studies with multiple strengths of identical or closely related formulation. The possible drawback of bracketing design is the uncertainty of similarity between the extremes and the intermediate levels. Both FDA and ICH suggest that the pharmaceutical industry gets a mutual agreement about applying bracketing design before the protocol in the stability study is executed.

An example of a bracketing design is given in Table 1 (ICH Q1D). This example is based on a product available in three strengths and three container sizes. It should be demonstrated that the 15-milliliter (ml) and 500-ml high-density polyethylene container seizes truly represent the extremes. The batches for each selected combination should be tested at each time point as in a full design.

Table 1. Example of Bracketing Design (ICH Q1D)

Strength		50mg			75mg			100mg		
Batch		1	2	3	1	2	3	1	2	3
Container Size	15 ml	T	T	T				T	T	T
	100 ml									
	500 ml	T	T	T				T	T	T

Key: T= Times when sample is tested e.g., T={0, 3, 6, 9, 12, 18, 24, 36}

Matrixing Design

As defined in the glossary of the parent guidance (ICH Q1D), matrixing is the design of a stability schedule such that a selected subset of the total number of possible samples for all factor combinations would be tested at a specified time point. At a subsequent time point, another subset of samples for all factor combinations would be tested. The design assumes that the stability of each subset of samples tested represents the stability of all samples at a given time point. The differences in the samples for the same drug product should be identified as, for example, covering different batches, different strengths, different sizes of the same container closure system, and possibly, in some cases, different container closure systems.

Matrixing design can be applied to the levels of the experimental factors, sampling times, and both levels of experimental factors and sampling times. Nordbrock suggested several matrixing designs for 2 years studies in 1992. According to Fairweather (1994),

factors such as strength and container size were acceptable by FDA for matrixing. The factors such as closer system, orientation of container during storage, packaging form, and batch size are possibly acceptable but the dosage form, manufacturing site, temperature, humidity, light and source of raw drug are unacceptable for matrixing.

An example of matrixing designs for a product with three strengths and three container sizes are given in Table 2 (ICH Q1D). Table 2-1 shows a design with matrixing on time points only and Table 2-2 depicts a design with matrixing on time points and factors. In Table 2-1, all combinations of batch, strength, and container size are tested, while in Table 2-2, certain combinations of batch, strength and container size are not tested.

Table 2. Examples of Matrixing Designs for a Product with Three Strengths and Three Container Sizes (ICH Q1D)

A. Matrixing on Time Points

Strength	S1			S2			S3		
Container Size	A	B	C	A	B	C	A	B	C
Batch 1	T1	T2	T3	T2	T3	T1	T3	T1	T2
Batch 2	T2	T3	T1	T3	T1	T2	T1	T2	T3
Batch 3	T3	T1	T2	T1	T2	T3	T2	T3	T1

B. Matrixing on Time Points and Factors

Strength	S1			S2			S3		
Container Size	A	B	C	A	B	C	A	B	C
Batch 1	T1	T2		T2		T1		T1	T2
Batch 2		T3	T1	T3	T1		T1		T3
Batch 3	T3		T2		T2	T3	T2	T3	

Key

Time-point	0	3	6	9	12	18	24	36
T1	X		X	X	X	X	X	X
T2	X	X		X	X		X	X
T3	X	X	X		X	X		X

S1, S2, and S3 are different strengths. A, B, and C are different container sizes.
X= Sample tested

Due to the reduced amount of data collected, matrixing design on factors other than time points generally has less precision in shelf life estimation and yields a shorter shelf life than the corresponding full design. In addition, such a matrixing design may have insufficient power to detect certain main or interaction effects, thus leading to incorrect pooling of data from different design factors during shelf life estimation. If there is an excessive reduction in the number of factor combinations tested and data from the tested factor combinations cannot be pooled to establish a single shelf life, it may be impossible to estimate the shelf lives for the missing factor combinations.

In 1996, Murphy introduced the uniform matrixing design by using the same test times for all the combinations of the design factors. The advantage of uniform matrixing design is the simplicity of study design and reduces the variability of the slope of the regression line. The disadvantage of uniform matrixing design is that there is no information obtained at the same times. The unavailability of the information may mislead the conclusion of stability of the drug product under investigation, especially significant confounding and interaction effect with sampling time points are present.

Statistical Methods

FDA Minimum Approach

The FDA stability guideline suggests that a drug expiration dating be determined as the time at which the 95% lower confidence bound of the mean degradation curve of the drug characteristic intersects the lower product specification limit, assuming that the drug characteristic (e.g., potency) decrease linearly with time. Thus, the following linear regression model is useful.

$$Y_j = \alpha + \beta X_j + e_j, \text{ j = 1,…., n.} \qquad (2\text{-}1)$$

where Y_j is the assay result of the potency (percent of labeled claim) at time X_j, α and β are the intercept and slope (i.e., stability loss over time), X_j's are the sampling time points selected in the stability study, e_j's are random errors in observing Y_j which are independent and identically distributed as N $(0, \sigma^2)$.

FDA guideline requires that at least three batches be tested to allow the estimate of batch-to-batch variability. This refers to the batch similarity, which can be examined by preliminary test to verify the equality of slope and intercepts of all batches. The guidelines also indicates that the preliminary tests for the equality of slopes and equality of intercepts be performed at 0.25 level of significance which was suggested by Bancroft (1964). If the preliminary tests for the equality of slopes and the equality of intercepts are not rejected at 0.25 level of significance, then the common slope and intercept are estimated on the pooled stability data from all batches. Note that as pointed out by Ruberg and Stegeman (1991), the use of the significance level of 25% for assessing batch-to-batch similarity for poolability (or combinability) might penalize good batches in drug shelf life estimation. On the other hand, if preliminary tests are rejected at the 0.25 level of significance, the minimum of expiration-dating periods obtained from individual batches is considered as the drug expiration-dating period. When a preliminary test for batch-to-batch variation is significant, the FDA suggests that the minimum of the individual estimated shelf lives obtained from individual batches be considered as the estimated shelf life of the drug product. The minimum approach for determination of shelf life, however, has received a considerable amount of criticism because it lacks of statistical justification (Chow and Shao, 1991) and suffers from some shortcomings (Ruberg and Stegeman, 1991). The advantages of the minimum approach is simple to

compute and relatively conservative. The shelf life only gives a guideline of how long a drug product should remain on the shelf, and this approach is easy to implement for the FDA regulations but it lacks statistical justifications because the minimum does not have a specific statistical meaning.

Parametric Approach

For a single batch, the FDA stability guideline indicates that an acceptable approach for drug characteristics that expected to decrease in potency with time is to determine the time at which the 95% one-sided lower confidence bound for the mean degradation curve interests the acceptable lower product specification limit. For multiple batches, if a preliminary test for batch-to-batch variation is not significant at the 0.25 level, then all batches are considered from the same population of production batches with a common degradation pattern, and a single estimated shelf life can be obtained by combining data from the different batches.

When there is no batch-to-batch variation, the model is followed,

$$Y_{ij} = \beta' X_{ij} + e_{ij}, \text{ i = 1,….., k, j = 1,….., n.}$$

Using the ordinary least squares method, we obtain the following 95% lower confidence bound for $\beta' x(t)$:

$$L(t) = \hat{\beta}' x(t) - t_{0.95;nk-p} \sqrt{x(t)'(X_1' X_1 + \cdots + X_k' X_k)^{-1} x(t) SSR/(nk-p)}$$

where $x(t)$ is x_{ij} with t_{ij} replaced by t and w_{ij} (package type or strength) fixed at a particular value, $\hat{\beta}$ is the ordinary least squares estimator of β, $X_i = (x_{i1},.....,x_{in})'$, SSR is the sum of squared residuals, and $t_{0.95;nk-p}$ is the 95th percentile of the t-distribution with nk-p degrees of freedom. According to the FDA stability guideline, the estimated shelf

life can be determined as $\hat{t}^* = \inf\{t : L(t) \le \eta\}$, where η is the given lower specification limit. If $L(t)$ is decreasing in t, $\hat{t}^*$ is the solution of $\eta = L(t)$, i.e., $\eta = L(\hat{t}^*)$. Since $\beta' x(t)$ is the average drug characteristic at time t, the true shelf life is $t^* = \inf\{t : \beta' x(t) \le \eta\}$. Note that t^* is nonrandom but is unknown. The estimated shelf life $\hat{t}^*$ is actually a 95% lower confidence bound for t^*, since $P(\hat{t}^* \rangle t^*) \le P(L(t^*) \rangle \eta) = P(L(t^*) \rangle \beta' x(t^*)) = 0.05$.

The minimum approach is conservative and does not consider the batch-to batch variation. Chow and Shao (1991) pointed out that the presence of batch-to batch variation has an impact on the determination of shelf life. The estimation procedures for shelf life with random batch effect are investigated by many researches.

Suppose that there are n batches in the stability analysis. Denote Y_{ij} the continuous outcome of the characteristic for the j^{th} measurement of the i^{th} batch i=1,2,…., n and j=1,2,……, m_i. Then in general the linear regression model assumes that

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + e_{ij}, \qquad (2\text{-}2)$$

where X_{ij} are the associated covariates, $(\alpha_i, \beta_i)^T$ are the parameters and e_{ij} are the independent random errors. The superscript T denotes the transpose of vector or matrix. When the batch-to-batch variation is of little concern, it is assumed that $\alpha_i = \alpha_0$ and $\beta_i = \beta_0$ for any i. The multiple linear regression models with the method of the ordinary least squares are suggested by Chow and Shao (2002) to obtain the estimates of $(\alpha_0$ and $\beta_0)^T$. The random effects models have been advocated to deal with the batch-to-batch variation.

Semiparametric Approach

The semiparametric regression approach doesn't need a distribution assumption for the error structure, thus it is more flexible than the usual parametric approach. In case that the error distribution does follow a Normal distribution, as assumed by the parametric approaches, the semiparametric approach may be less efficient compared to the parametric approaches in terms of computing and variance estimating. But in practice, the normality assumption isn't always met, as pointed in Chen, Pong and Xing (2003). Specifically, the following semiparametric model is assumed:

$$h(Y_{ij}) = \alpha_i + \beta_i X_{ij} + e_{ij}, \qquad (2\text{-}3)$$

where e_{ij} are independent zero-mean deviates with unknown density function of $f(\cdot)$ and $h(\cdot)$ is some known monotonic transformation function. The model above is quite flexible without assuming the underlying distribution of normality or constant variance and also it does not necessarily assume the parametric distribution on $(\alpha_i, \beta_i)^T$ as random effects but leaves them as fixed effects, when appropriate rank-based inference procedures are utilized for the corresponding models (Chen, Pong and Xing, 2003; Jung and Ying, 2003; Koul, Sievers and Mckean, 1987; Rashid 2003). The model (2-3) becomes (2-1) for the stability analysis when one batch of stability data is to be analyzed, that is, n = 1. So the index i is ignored to simplify the notations and $h(\cdot)$ is an identity link without loss of generality. Under the null hypothesis of $\beta = 0$, a rank test statistic is

$$U = \sum_{j=1}^{m} \left(X_j - \bar{X} \right) R\left(Y_j\right),$$

where $\overline{X}$ is the mean of X's and $R(Y_1)$, $R(Y_2)$,……, $R(Y_m)$ are the ranks of Y_1, Y_2,….., Y_m respectively. When β is not necessarily zero; a straightforward extension is to consider the residuals of $e_j = Y_j - (\alpha + \beta X_j)$:

$$U(\beta) = \sum_{j=}^{m} \left(X_j - \overline{X}\right) R\left(e_j\right).$$

Apparently $E[U(\beta);\beta] = 0$ for any $\beta \in B \subset R$, where B is the parameter space and R is the real line. It is reasonable to solve the unbiased estimating equation of U(β) to obtain appropriate estimates of β;

$$U(\hat{\beta}) = 0$$

CHAPTER 3
STABILITY ANALYSIS

This chapter shows the pair-wise estimation that is proposed as a new method and the proof for the comparability of random errors, the computing method for estimating the parameter values of the model, and shelf life determination.

Pair-wise Estimation

The following is a simple linear regression model:

$$Y_i = \alpha + \beta X_i + e_i, \quad i = 1,2, \ldots, n \qquad (3\text{-}1)$$

where Y_i are dependent variables, X_i are independent variables, and e_i are mean 0 random errors.

The comparability of random errors is proved in the next step. When the random errors are random, any pair of $\{e_1, e_2, e_3, \ldots\ldots, e_n\}$ are comparable, that is to say,

$$E\{I[(Y_i - \beta X_i) \le (Y_j - \beta X_j)]\} = E\{I(e_i \le e_j)\} = 1/2 \qquad (3\text{-}2)$$

Here is the proof:

- Order the e_i's, thus we have $\{e_{(1)}, e_{(2)}, \ldots\ldots, e_{(n)}\}$, where $e_{(i)}$ is the i^{th} order statistics.
- The total number of pairs of comparison including the case where i=j is n^2 and the number of pairs of self-comparison is n. Thus the number of pair comparison excluding self-comparison is n^2-n.
- For fixed i,
  $I(e_{(i)} \le e_{(j)}) = I(i \le j)$
  $$\sum_{j=1, j\neq i}^{n} I(e_{(i)} \le e_{(j)}) = \sum_{j=1, j\neq i}^{n} I(i \le j) = n - i$$

$$\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} I(e_{(i)} \le e_{(j)}) = \sum_{i=1}^{n}(n-i) = n^2 - \sum_{i=1}^{n} i = (n^2-n)/2$$

- Since $I(e_{(i)} \le e_{(j)})$ is a 1/0 variable, $\Pr\{e_{(i)} \le e_{(j)}\} = 1/2$ that means the density distribution for $e_{(i)}$ and $e_{(j)}$ is symmetric, and also $e_{(i)}$ and $e_{(j)}$ are independent. This is proved by the definition 4.2.5 for two independent random variables in the book, "Statistical Inference", Casella and Berger, 2002 (p152).

The proposed pair-wise estimating equation is following:

$$E\{I(e_i \le e_j)\} = 1/2$$
$$\Pr(e_i \le e_j) = \Pr(e_i - e_j \le 0) = 1/2$$
$$\Pr(Y_i - Y_j - \beta(X_i - X_j) \le 0) = 1/2$$

The proposed estimating equation for a general β is:

$$U(\beta) = \sum_{i=1}^{n}\sum_{j=i+1}^{n}\{I[(Y_i - \beta X_i) \le (Y_j - \beta X_j)] - 1/2\}$$
$$= \sum_{i=1}^{n}\sum_{j=i+1}^{n} I[Y_i - Y_j - \beta(X_i - X_j) \le 0] - (n^2-n)/4 = 0 \qquad (3\text{-}3)$$

Apparently, E(U(β)) = 0 for any β in the parameter space, $\hat{\beta}$ is obtained by solving the unbiased estimating equation (3-3). The solution of β could be defined to be $\hat{\beta}$ that satisfies U(β+)U(β-)≤ 0, if there is no exact β that allows U(β)=0.

α is not estimable from the estimating equation since the rank won't change by a constant factor. In the former papers, median of $Y_i - \hat{\beta}X_i$ or Walsh average $(e_i + e_j)/2$ under the symmetry distribution are used to determine α (Hettmansperger, 1984). We proposed to use the median of $Y_i - \hat{\beta}X_i$ as $\hat{\alpha}$.

The feature of the asymptotic distribution of U(β) is following:

- E(U(β)) = 0 and U($\hat{\beta}$) = 0
- The variance formula for $\hat{\beta}$, where U($\hat{\beta}$) = 0

$$Var(U(\beta)) = \sum_{j=1}^{m}\sum_{k=j+1}^{m} Var[I(Y_{ij} - \beta X_{ij} \rangle Y_{ik} - \beta X_{ik})]$$

$$+ \sum_{j=1}^{m}\sum_{k=j+1}^{m}\sum_{j^*=1}^{m}\sum_{k^*=j^*+1}^{m} Cov[I(Y_{ij} - \beta X_{ij} \rangle Y_{ik} - \beta X_{ik}), I(Y_{ij}^* - \beta X_{ij}^* \rangle Y_{ik}^* - \beta X_{ik}^*)]$$

$$\cong (n^2\text{-}n)/8 + \{(n\text{-}1)(n\text{-}2)(2n\text{-}3)/6 + (n\text{-}1)(n\text{-}2)/2\}/24 + \{(n^3\text{-}2n^2\text{+}n)/2$$

$$- n(n\text{-}1)(2n\text{+}1)/3\}/12$$

- When the sample size n is large, it follows from the Central Limit Theorem:
  $U(\beta)/\sqrt{Var(u(\beta)} \sim N(0,1)$

Batch-to-batch Variation Models

The model for stability data is following:

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + e_{ij}, \quad i = 1,2, \ldots, n, j = 1,2, \ldots., mi \qquad (3\text{-}4)$$

where Y_{ij} are the continuous outcome of the characteristics for the j^{th} measurement of the i^{th} batch, X_{ij} are the associated covariates, α_i andβ_i are the parameters and e_{ij} are the independent random errors.

When the batch-to-batch variation is of concern, three cases are considered.

- Slopes are the same, but intercepts are different, i.e., $\beta_i = \beta$.
  The model and the estimating equation are:
  $Y_{ij} = \alpha_i + \beta X_{ij} + e_{ij}$, i = 1, 2, …, n, j = 1, 2, …, m_i.

  $$U(\beta) = \sum_{i=1}^{n} U_i(\beta)$$

  $$= \sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=j+1}^{m_i} I[(Y_{ij} - \beta X_{ij}) \le (Y_{ik} - \beta X_{ik})] - \sum_{i=1}^{n}(m_i^2 - m_i)/4 = 0$$

  For batch i, $\hat{\alpha}_i = median(Y_{ij} - \hat{\beta} X_{ij})$, j = 1, 2, …, m_i
  In case of the balanced data, i.e., $m_i = m$, i = 1, 2, …, n, then

  $$U(\beta) = \sum_{i=1}^{n} U_i(\beta)$$

  $$= \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=j+1}^{m} I[(Y_{ij} - \beta X_{ij}) \le (Y_{ik} - \beta X_{ik})] - n*(m_i^2 - m_i)/4 = 0$$

- Slopes are different, but intercepts are the same, i.e., $\alpha_i = \alpha$.

The model is:

$Y_{ij} = \alpha + \beta_i X_{ij} + e_{ij}$, i = 1, 2, …, n, j = 1, 2, …, m_i

For batch i, the estimating equation is:

$$U_i(\beta) = \sum_{j=1}^{m_i} \sum_{k=j+1}^{m_i} \{I[(Y_{ij} - \beta_i X_{ij}) \le (Y_{ik} - \beta_i X_{ik})] - 1/2\}$$

$$= \sum_{j=1}^{m_i} \sum_{k=j+1}^{m_i} I[Y_{ij} - Y_{ik} - \beta_i (X_{ij} - X_{ik}) \le 0] - (m_i^2 - m_i)/4 = 0$$

$\hat{\alpha} = median(Y_{ij} - \hat{\beta} X_{ij})$, i = 1, 2, …, n, j = 1, 2, …, m_i.

- Both slopes and intercepts are different.

  The model is:

  $Y_{ij} = \alpha_i + \beta_i X_{ij} + e_{ij}$, i = 1, 2, …, n, j = 1, 2, …, m_i

  For batch i, the pair-wise estimating equation is:

$$U_i(\beta) = \sum_{j=1}^{m_i} \sum_{k=j+1}^{m_i} \{I[(Y_{ij} - \beta_i X_{ij}) \le (Y_{ik} - \beta_i X_{ik})] - 1/2\}$$

$$= \sum_{j=1}^{m_i} \sum_{k=j+1}^{m_i} I[Y_{ij} - Y_{ik} - \beta_i (X_{ij} - X_{ik}) \le 0] - (m_i^2 - m_i)/4 = 0$$

$\hat{\alpha}_i = median(Y_{ij} - \hat{\beta} X_{ij})$, j = 1, 2, …, m_i

Shelf Life Determination

Bootstrapping is used to obtain the confidence interval of $\hat{\alpha} + \hat{\beta} X$. The Bootstrap method is an approach to statistical inference that makes few assumptions about the underlying probability distribution that describes the data. This approach assumes that empirical cumulative distribution function is a reasonable estimate of the unknown, population cumulative distribution function. Using the data as an approximation to the population density function, data is re-sampled with replacement from the observed sample to create an empirical sampling distribution for the test statistic under consideration. The empirical distribution of these calculated statistics is referred to as an empirical sampling distribution. This empirical sampling distribution can be used as an approximation to the theoretical population sampling distribution. To calculate scores that correspond the 2.5^{th} and 97.5^{th} confidence intervals, we simply find the scores that

correspond to the upper and lower percentiles of the distribution. For example, for 1000 bootstrap samples of the sample mean, we can calculate confidence intervals by finding the upper and lower 2.5^{th} and 97.5^{th} percentiles using the following approach: round[(0.05/2)*1000] = 25 for lower percentiles and round[{1-(0.05/2)}*1000] = 975. Moreover, the standard deviation of the bootstrapped test statistics is an approximation of the standard error of the test statistic under consideration. This approach to calculating bootstrap confidence intervals is called the Percentile Bootstrap method (Davison and Kuonen, Stat. Comp. & Stat. Graph. Newsletter; Herrington, 2004).

The Percentile Bootstrap method is utilized for determining the shelf life. The steps are as following:

- For a batch i, m_i pairs of (X_j, Y_j) are picked randomly with replacement.
- The pair-wise estimating equation is used to obtain $\hat{\beta}$
- Repeat the above steps with B times, thus we have $\hat{\beta}_b$ for b = 1, 2, …, B.
- Take the 2.5 percentile of $\hat{\beta}_b$ as the lower confidence bound for $\hat{\beta}$.
- $\hat{\alpha}$ is estimated as the median of $Y_{ij} - \hat{\beta}_{LB} X_{ij}$.
- For a continuous X^* including the range of X, the estimate of $\hat{Y}^*$ is computed by $\hat{Y}^* = \hat{\alpha} + \hat{\beta}_{LB} X^*$.
- The time at which $\hat{Y}^*$ intersects the pre-specified limit is the shelf life.

CHAPTER 4
NUMERICAL STUDIES

Simulation

In our simulations, we study the performance of pair-wise estimation procedure under different error structures. For simplicity, we only consider the following model:

$Y_{ij} = \beta_1 X_{1i} + \beta_2 X_{2ij} + e_{ij}$, for i = 1, 2, …, n and j = 1, 2, …, m,

where Y_{ij} is the characteristic of interest of the j^{th} measurement for i^{th} batch, X_{1i} is covariate independent of j (e.g. gender), X_{2ij} is the time of the j^{th} measurement for the i^{th} batch (e.g. seasonality), e_{ij} is random error of the j^{th} measurement for the i^{th} batch, β_1 is constant and β_2 is slope. β_1 and β_2 are fixed to be1. This is the common slope model, but it is straightforward to expand all the simulations described in chapter 3 to other occasions with batch-to-batch variation and simple liner regression model.

An example of simulation is showed to illustrate how to solve for β using the proposed estimating equation:

- A vector of random error is generated using the Normal/Uniform/Lognormal distribution.
- β is fixed to be 1.
- X is a vector of {1, 2, …, 100}.
- Y_i, for i = 1, 2, …., 10 is generated as $Y_i = \beta X_i + e_i$.
- The pair-wise estimation method is used to get $\hat{\beta}$.

The U(β) is obtained by the pair-wise estimation and it is plotted against β, the value of β at which U(β) intersects the x-axis is $\hat{\beta}$ as desired. Figure 2 shows that the estimated value of β is one.

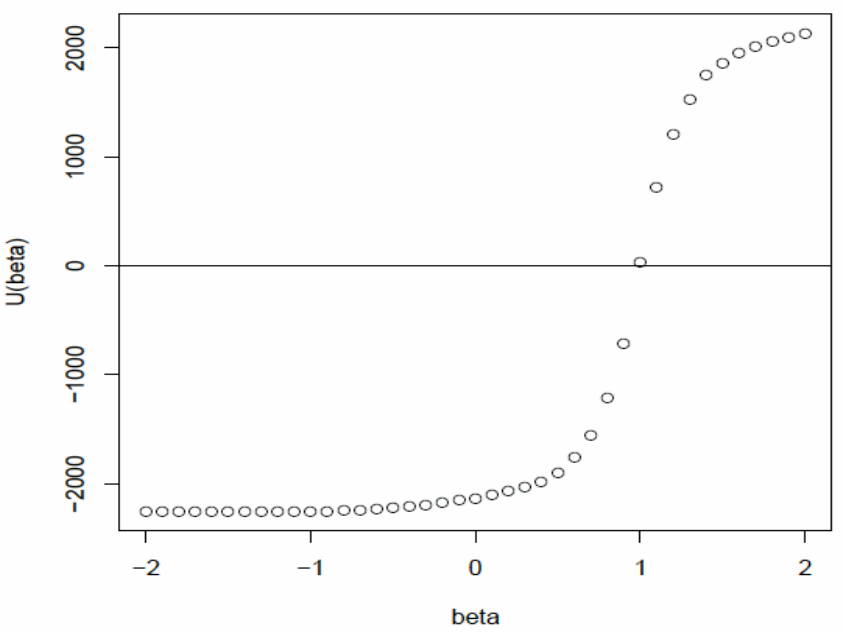

Figure 2. Pair-wise Estimating Equation

The FDA (1987) recommends at least three batches be tested to obtain the shelf life for a new drug application (NDA) submission. In our simulation study, we considered n = 1, n = 3 and n = 5. Furthermore, we assume that the characteristic of interest increase over time, as β = 1. Three error distributions are considered: (1) Normal [$e_{ij} \sim N(0,1)$]; (2) Uniform [$e_{ij} \sim U(-1, 1)$]; and (3) Log-normal [Log $e_{ij} \sim N(0,1)$]. One thousand sets of data are provided for each error distribution and for the number of batch n = 1, n = 3 and n = 5 respectively. The estimated values of β and their standard deviations are reported in Table 3. Histograms of the estimated slopes are shown in Fig. 3. The standard deviation decreases as the number of batches increases. When we use U(-1,1) for the error distribution, it has the smallest standard deviation.

Table 3. Mean and Standard Deviation of Estimated Model Parameter β

Parameter	# of batches	Error distribution		
		$e_{ij} \sim N(0,1)$	$e_{ij} \sim U(-1,1)$	Log $e_{ij} \sim N(0,1)$
β	n = 1	0.9912(0.11)	0.9990(0.07)	0.9995(0.12)
	n = 3	0.9970(0.07)	1.0018(0.04)	0.9992(0.06)
	n = 5	1.0016(0.05)	0.9992(0.03)	0.9979(0.04)

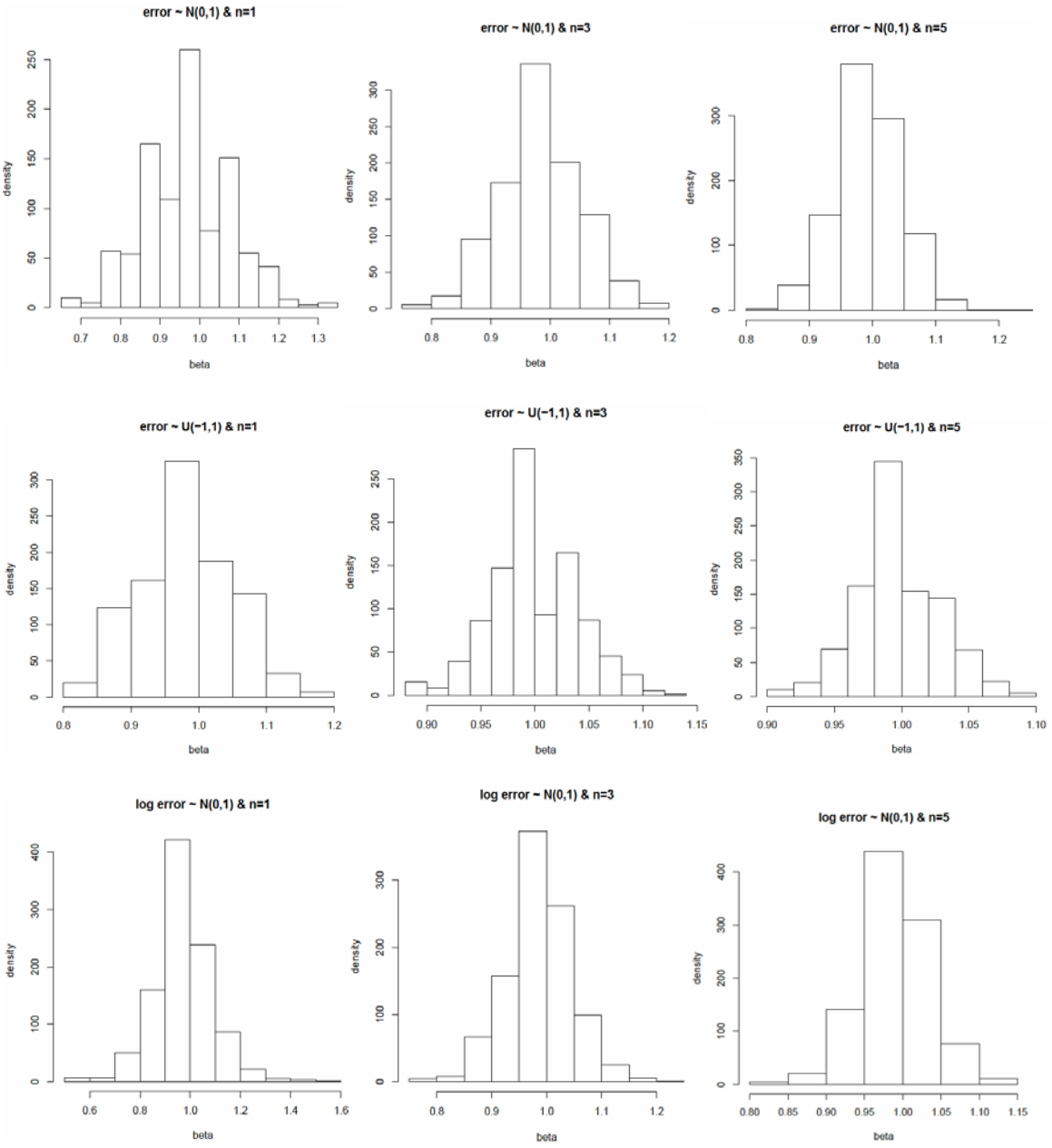

Figure 3. Histograms of Estimated Slopes. Note that true slope is β = 1.

Analysis of Shao and Chow's Data

In Shao and Chow (1994), a data set of the stability study on 300mg tablets of a drug product to determine appropriate labeled shelf life was given and analyzed. The tablets from five batches were stored at room temperature in two types of containers: high -density polyethylene bottle and blister packages. Tests of potency were conducted at 0, 3, 6, 9, 12 and 18 months. The data are plotted in Fig. 4. The potency is generally degrading as time progresses.

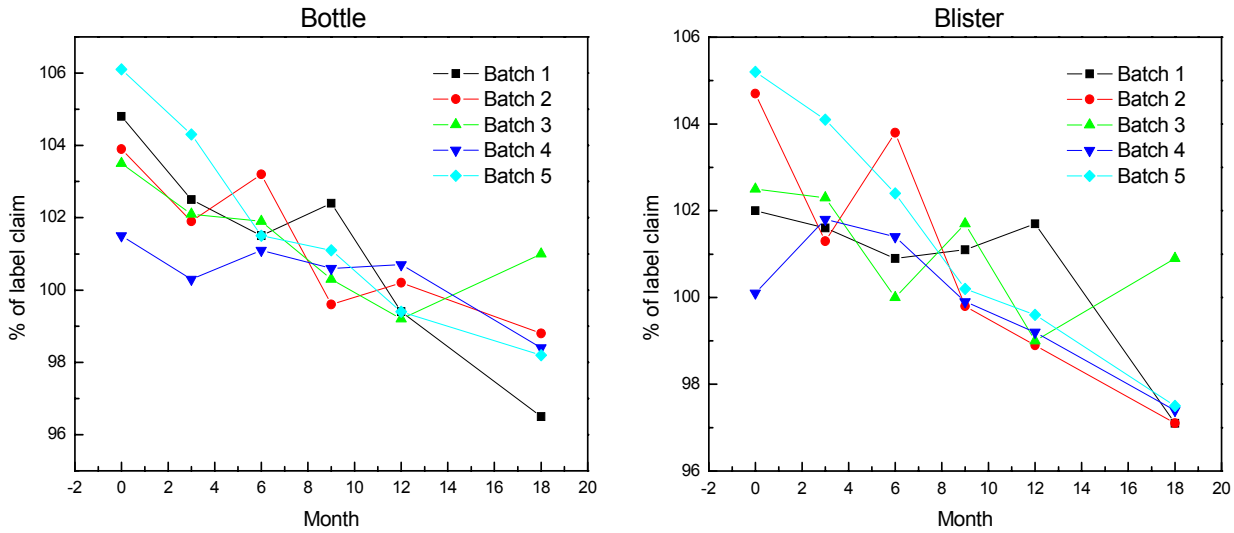

Figure 4. Stability Study Result of a 300mg tablet. Source: Shao and Chow' Data (1994)

A closer look at Fig. 4 would find that the variation of degradation curves tends to be bigger in intercepts while the slopes of these curves tend to be similar. Therefore, the model with common slope and different intercepts:

$$Y_{ij} = \alpha_i + \beta X_{ij} + e_{ij}$$

is fitted to the data. The estimating function is non-increasing with β and the estimates of β are –0.330 and –0.293 for the HDPE bottle container and blister package, respectively.

Furthermore, simple liner regression models fitted with the OLS for the HDPE bottle container and blister package, respectively and the estimates of β are –0.289 and –0.278 for each case. For the HDPE bottle container, the estimates of intercept for five batches are 103.485, 104.030, 103.385, 103.325, and 104.105. For the blister package, the estimates of the intercept for five batches are then 102.5685, 102.4265, 102.8475, 102.6765, and 103.6370.

While the model with different slopes and intercepts:

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + e_{ij}$$

is fitted to the Shao and Chow's data as well. The estimates of (α_i, β_i)s are showed in Table. 4.

Table 4. Estimates of α and β under the Model with Different Slopes and Intercepts

Container	Estimates	Batch				
		1	2	3	4	5
Bottle	α	104.55	103.10	103.32	101.62	105.77
	β	-0.44	-0.24	-0.32	-0.13	-0.51
Blister	α	101.97	102.59	102.51	102.39	105.15
	β	-0.11	-0.31	-0.09	-0.27	-0.45

The estimated mean degradation curves form both models for their respective container type are also plotted in Fig. 5.

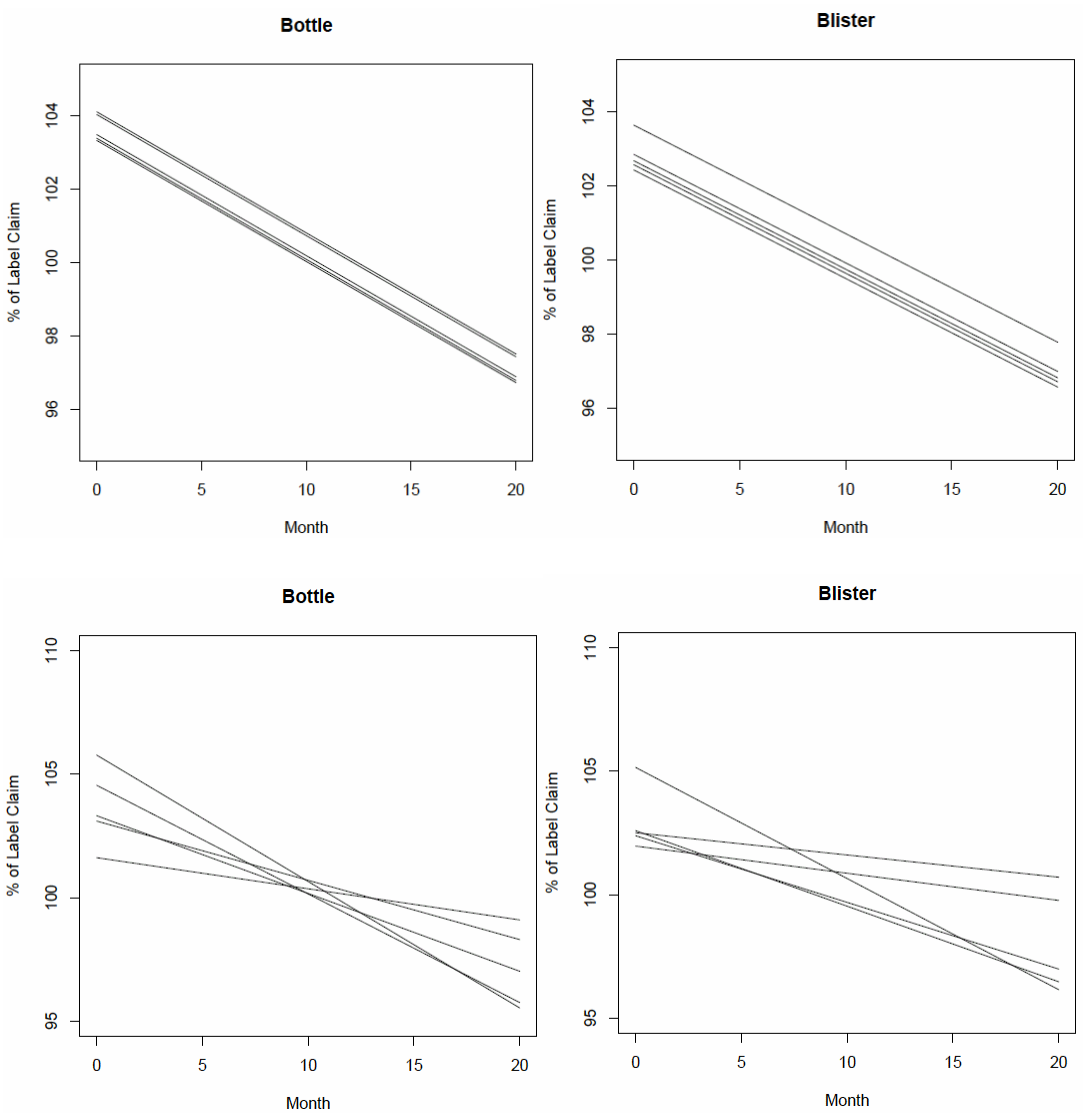

Figure 5. Estimated Mean Degradation Curves of Shao and Chow's Data (The plots for the model with common slope and different intercepts are in the first row and the plots for the model with different slopes and intercepts are in the second row.)

Assuming batches with bottle container have the same slope and so do the batches in blister, the Percentile Bootstrap method is used to compute the shelf life for bottle and blister container. The shelf life for bottle container is 24 months, and 22 months for blister with the specification limit of 90%. The plot of the shelf life for the HDPE bottle container and blister package is shown in Figure 6.

The shelf life by the pair-wise estimation method in this paper is comparable with Shao and Chow (1994)'s, and Chen, Pong and Xing (2003)'s method. The comparison of shelf life with other approaches is illustrated in Table 5. The shelf lives for bottle and blister are longer than those of other approaches. The reason is that we used a Bootstrap method to compute shelf life, which approximate the asymptotic distribution of the lower confidence interval of $\hat{Y} = \hat{\alpha} + \hat{\beta}X$. If the variance of the joint distribution of α and β were obtained, the shelf lives from our approach would be more exact.

Table 5. Comparison of Shelf Life (month) of Shao and Chow's Data from Other Approaches

Method	Shelf life	
	Bottle	Blister
FDA Minimum approach	26	26
Shao and Chow's approach with ε = 0.04	21	21
Shao and Chow's approach with ε = 0.05	22	21
Chen, Pong and Xing's approach	22	21
Pair-wise Estimation method	24	22

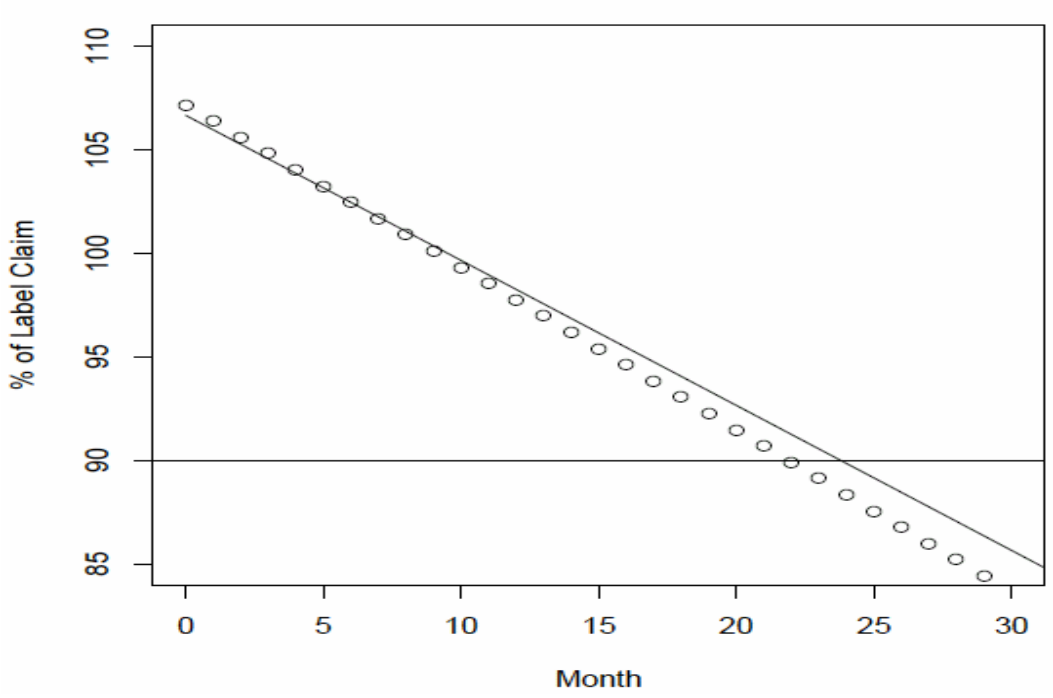

Figure 6. Shelf Lives of Shao and Chow's Data (bottle-line; blister-circle)

CHAPTER 5
COMPARISON OF STABILITY STUDY DESIGNS

The design of a stability study is intended to establish a shelf life applicable to all future batches of the drug substance or product manufactured under similar circumstances based on testing a limited number of batches of a drug substance or product. The stability study should be well designed so the shelf life of the product can be estimated with the high degree of accuracy and precision.

Several different types of design can be applied to the stability study. A full design is one in which samples for every combination of all design factors are tested at all time points, while a reduced design is one which in samples for every factor combination are not all tested at all time points. A reduced design can be suitable alternative to a full design when multiple design factors are involved in the product being evaluated, but the application of a reduced design has to be carefully assessed, taking into consideration any risk to the ability of estimating an accurate and precise shelf life. Bracketing and matrixing designs are two most commonly used reduced designs (Lin and Chen, 2003). The specific explanation about bracketing and matrixing designs was given in Chapter 2, literature review. We want to apply pair-wise estimation method to two reduced designs and compare a full design with a reduced design in this chapter.

Stability Analysis of Tsong, Chen and Chen's Data

To compare the three designs (Full design; Bracketing; Matrixing) for one component, an experimental factor with 3 levels, and 3 batches at each level as recommended by FDA are considered, assuming that there is no slope difference between

batches at the same factor level. The sampling time points used for the three designs are shown in Table 6.

Table 6. Sampling Time Points (month) for Three Designs

Container Size	3			30			100		
Batch	1	2	3	1	2	3	1	2	3
Full	T_0	T_0	T_0	T_0	T_0	T_0	T_0	T_0	T_0
Bracketing	T_0	T_0	T_0				T_0	T_0	T_0
Matrixing	T_1	T_2	T_3	T_2	T_3	T_1	T_3	T_1	T_2

T_0 = (0, 3, 6, 9, 12, 18); T_1 = (0, 6, 9, 12, 18); T_2 = (0, 3, 9, 12, 18); T_3 = (0, 3, 6,12,18)

Tsong, Chen and Chen's data (2003) are used to compare three designs as a data example. A stability design with one levels of strength (10 mg), three levels of container sizes (3, 30 and 100 tablets) and three batches is considered. Measurements were made at 0, 3, 6, 9, 12 and 18 months.

The parameters, i.e., slope and intercept, and shelf life can be estimated by the pair-wise estimation method proposed in this thesis. The models used are batch-to-batch variation in different slopes and different intercepts, and the occasion of the same slope and different intercepts. The models are following:

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + e_{ij}, \quad \text{i = 1, 2, 3, and j = 1, 2, 3, 4, 5, 6/ 1, 2, 3, 4, 5}$$

$$Y_{ij} = \alpha_i + \beta X_{ij} + e_{ij}, \quad \text{i = 1, 2, 3, and j = 1, 2, 3, 4, 5, 6/ 1, 2, 3, 4, 5}$$

Bootstrap method is used to determine the shelf lives finally as it is given in Chapter 3.

Example of Full Design

A full design is also referred to as a complete factorial design, since samples for every combination of all design factors, i.e., batch and container size, are tested at all time points. The example of full design is given in Table 7. The total number of samples tested is N = 3 * 3* 6 = 54. The estimated values of parameters are shown in Table 8.

Table 7. Assay Measurements of Full Design from Tsong, Chen and Chen's Data (2003)

		Time in Months					
Container Size	Batch	0	3	6	9	12	18
3	1	100	101	101	100	99	98
	2	101	100	101	99	98	99
	3	100	101	100	101	99	99
30	1	101	100	99	99	98	97
	2	102	101	99	100	99	98
	3	101	101	99	99	99	97
100	1	102	101	100	99	100	99
	2	102	101	101	99	99	98
	3	101	100	99	100	99	98

Table 8. Parameter Estimates of Tsong, Chen and Chen's Full Design Data

Design	Container Size	Batch	Batch-to-batch Variation			
			$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\alpha}_i$	$\hat{\beta}$
Full	3	1	101.71	-0.21	101.60	-0.20
		2	100.90	-0.20	100.90	
		3	101.50	-0.20	101.50	
	30	1	100.81	-0.22	100.81	-0.22
		2	101.73	-0.21	101.81	
		3	100.90	-0.20	100.99	
	100	1	101.80	-0.20	101.79	-0.19
		2	101.94	-0.24	101.50	
		3	101.13	-0.15	101.50	

Example of Bracketing Design

The design of a stability schedule such that only samples on the extremes of certain design factors are tested at all time points as in a full design. The design assumes that the stability of any intermediate levels is represented by the stability of the extremes tested. The following table shows the example of bracketing design. Bracketing is applicable if the strengths are identical or very closely related in composition, when we want to test the drug product's strength and also it can be applied to different container sizes or different fills in the same container closure system.

Table 9. Assay Measurements of Bracketing Design from Tsong, Chen and Chen's Data

		Time in Months					
Container Size	Batch	0	3	6	9	12	18
3	1	100	101	101	100	99	98
	2	101	100	101	99	98	99
	3	100	101	100	101	99	99
30	1	X	X	X	X	X	X
	2	X	X	X	X	X	X
	3	X	X	X	X	X	X
100	1	102	101	100	99	100	99
	2	102	101	101	99	99	98
	3	101	100	99	100	99	98

X means "NO" experimental measurement.

The result of stability analysis is the same with that of full design, but there is no analysis result of container size 30 since we assume that container sizes 3 and 100 are tested.

Example of Matrixing Design

The design of a stability schedule such that a selected subset of the total number of possible samples for all factor combinations is tested at a specified time point. At a subsequent time point, another subset of samples for all factor combinations is tested. The following Table 10 shows the example of matrixing design and the estimates of parameters by pair-wise estimation are given in Table 11.

Table 10. Assay Measurements of Matrixing Design from Tsong, Chen and Chen's Data

		Time in Months					
Container Size	Batch	0	3	6	9	12	18
3	1	100	X	101	100	99	98
	2	101	100	X	99	98	99
	3	100	101	100	X	99	99
30	1	101	100	X	99	98	97
	2	102	101	99	X	99	98
	3	101	X	99	99	99	97
100	1	102	101	100	X	100	99
	2	102	X	101	99	99	98
	3	101	100	X	100	99	98

The matrixing design assumes that the stability of each subset of samples tested represents the stability of all samples at a given time point.

Table 11. Parameter Estimates of Tsong, Chen and Chen's Matrixing Design Data

Design	Container Size	Batch	Batch-to-batch Variation			
			$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\alpha}_i$	$\hat{\beta}$
Matrixing	3	1	101.50	-0.20	101.6	-0.21
		2	100.95	-0.21	100.8	
		3	101.48	-0.22	101.4	
	30	1	101.00	-0.23	101.00	-0.23
		2	101.81	-0.22	101.88	
		3	100.90	-0.20	101.04	
	100	1	102.00	-0.20	102.00	-0.20
		2	102.00	-0.23	101.80	
		3	101.26	-0.21	101.20	

Shelf Life Comparison of Three Designs

The shelf lives of full, bracketing, and matrixing designs are obtained the pair-wise estimation method and bootstrap and the result of Tsong, Chen and Chen's data analysis is given in Table 12 and Figure 7. Note that the specification limit is fixed at 95% of Label Claim.

Table 12. Shelf Lives (months) of Tsong, Chen and Chen's Three Designs

Design	Container Size	Shelf Life
Full	3	26
	30	22
	100	25
Bracketing	3	26
	100	25
Matrixing	3	25
	30	22
	100	27

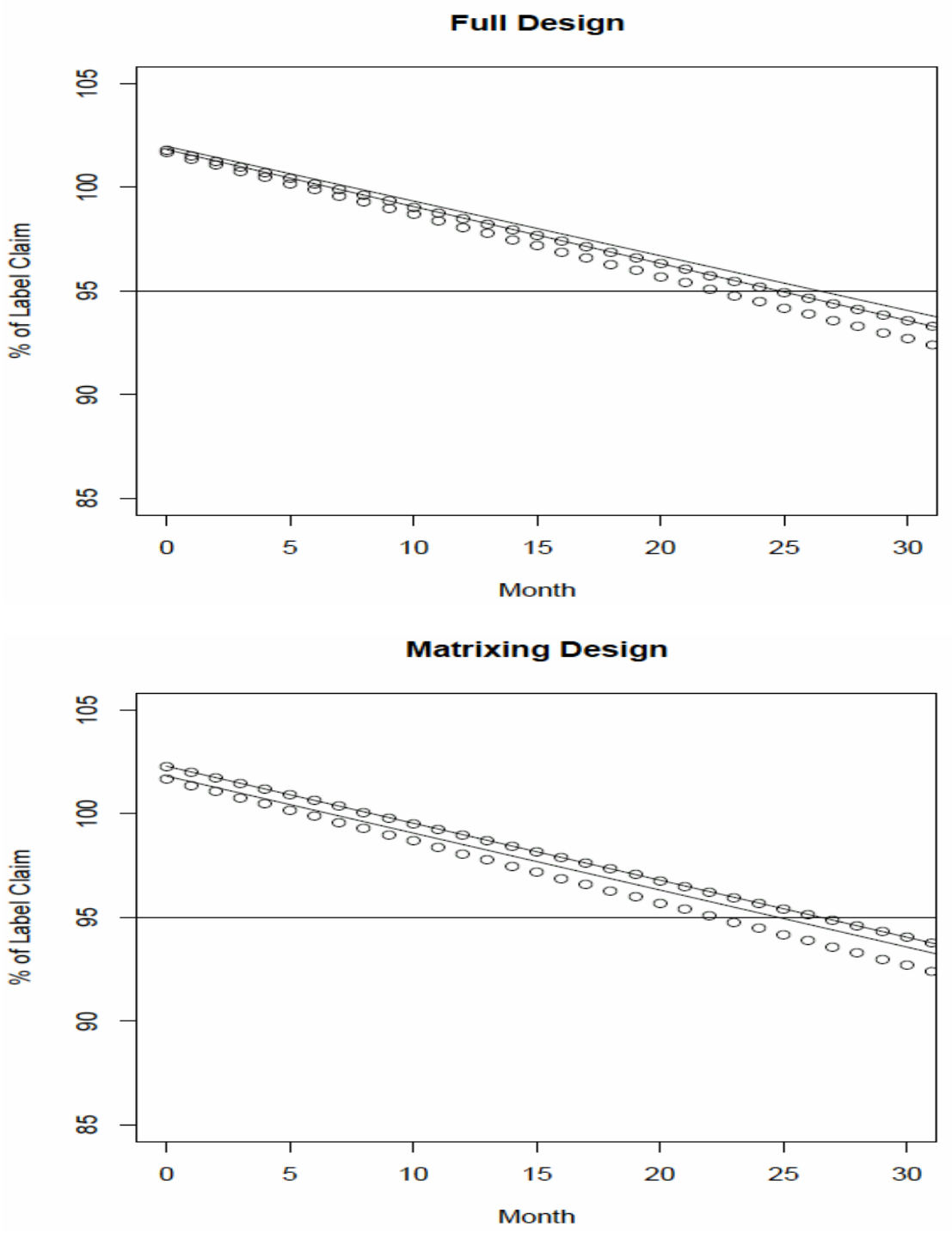

Figure 7. Shelf Lives of Tsong, Chen and Chen's Data (–: 3 tablets, O: 30 tablets and ∅: 100 tablets)

CHAPTER 6
FUTURE RESEARCH

The pair-wise estimation approach based on the semiparametric rank regression is proposed in this thesis. This approach does not need the assumption for the error distribution, thus it is more flexible than the usual parametric approach, but there is the disadvantage that it is less efficient than the parametric approach with an aspect of computation and variance estimation.

One of methods to estimate α is introduced in Chapter 3, which is to use the median of $Y_i - \hat{\beta}X_i$. The intercept term α is not estimable by the pair-wise estimation because the rank is invariant to a constant. That's why we proposed the median estimation in this thesis. The Walsh average could be applied too.

We proposed the Bootstrap method to determine the shelf life, but it is possible that the value estimated is rough. The more exact value could be obtained by the theoretical 95% confidence interval:

$$[\hat{\alpha} + \hat{\beta}X_i - Z_{0.025}\sqrt{(1, X_i)\hat{V}(1, X_i)^T}, \hat{\alpha} + \hat{\beta}X_i + Z_{0.025}\sqrt{(1, X_i)\hat{V}(1, X_i)^T}]$$

where $\hat{V}$ is the variance-covariance matrix V of the joint distribution of α and β, and T denotes transpose.

There are some extensions of the proposed method and further investigations on the stability analysis of the drug product. Possible future research outlines are listed as follows:

1. We analyzed the stability study for the drug products with single component as data examples in Chapter 4 and 5. The pair-wise estimation can be applied to the stability study analysis for the drug products with even multiple components.

2. In case of stability study for multiple components, the components may not be independent, thus it is useful to take the interaction among components into determination of shelf life.

3. The approved specification limit is an important factor when we determine the shelf life too. The right selection of the specification limit on different drug products and components should be investigated.

4. We compared the full design with the reduced design including bracketing and matrixing designs of single components in Chapter5, but there is a limitation to analyze the reduced design, since sample points are sparse. Therefore, the method on how to analyze the reduced design more precisely needs to be added and further studied.

As we can see above, there are lots of subjects to solve in the stability study of drug products. Particularly, shelf life estimation for drug products with multiple components is one of the most important tasks to biostatistician since it includes complicated factors. Shelf life estimation is related not only to our health but also to the development of drug products directly, thus it should be further investigated and studied.

LIST OF REFERENCES

Bancroft, T. A. 1964. Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. Biometrics 20: 427-442.

Casella, George and Roger L. Berger. 2002. Statistical Inference. Pacific Grove, CA: Duxbury.

Chen, Ying Qing, Annpey Pong and Biao Xing. 2003. Rank regression in stability analysis. Journal of Biopharmaceutical Statistics 13: 463-479.

Chow, Shein-Chung and Annpey Pong. 1995. Current issues in regulatory requirements of drug stability. Journal of Food and Drug Analysis 3(2): 75-85.

Chow, Shein-Chung and Jun Shao. 1991. Estimating drug shelf-life with random batches. Biometrics 47: 1071-1079.

Chow, Shein-Chung and Jun Shao. 2001. Stability analysis with discrete response. Currently under revision for Journal of Biopharmaceutical Statistics.

Chow, Shein-Chung and Jun Shao. 2002. Statistics in Drug Research: Methodologies and Recent Developments. New York: Marcel Dekker.

Davison, A. C. and Diego Kuonen. Topics in statistical computing: An introduction to the Bootstrap with applications in R. Statistical Computing & Statistical Grapics Newsletter 13: 6-11.

Fairweather, W. R. 1994. Design of stability studies: FDA statistical perspective. Paper presented at symposium, American Association of Pharmaceutical Scientists Workshop on Stability Guidelines for Testing Pharmaceutical Products, Arlington, VA.

Food and Drug Administration/Center for Drug Evaluation and Research. 2002. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. Rockville: FDA.

Food and Drug Administration/Center for Drug Evaluation and Research. 1997. Guidance for Industry. Rockville: FDA.

Herrington, Rich, An introduction to the percentile Bootstrap using GNU S. http://www.unt.edu/rss/benchmarks/Rich/rss_matters_aug2001, last accessed May 14, 2004

Hettmansperger, Thomas P. 1984. Statistical Inferences Based on Ranks. New York: John Wiley.

Jung, Sin-Ho and Zhiliang Ying. 2003. Rank-based regression with repeated measurements data. Biometrika 90: 732-740.

Koul, Hira L., Gerald L. Sievers and Joseph Mckean. 1987. An estimator of the scale parameter for the rank analysis of linear models under general score functions. Scand J Statist 14: 131-141

Lin, Tsae-Yun Daphne and Chi Wan Chen. 2003. Overview of stability study designs. Journal of Biopharmaceutical Statistics 13: 337-354.

Murphy, J. R. 1996. Uniform matrix stability study designs. Journal of Biopharmaceutical Statistics. 6: 477-494.

Pong, Annpey. 2001. Comparing designs for stability studies and shelf life estimation for drug product with multiple components. Dissertation at the Temple University.

Rashid, M. Mushfiqur. 2003. Rank-based tests for non-inferiority and equivalence hypotheses in multi-centre clinical trials using mixed models. Statistics in Medicine 22: 291-311.

Ruberg, Stephen J. and James W. Stegeman. 1991. Pooling data for stability studies: Testing the equality of batch degradation slopes. Biometrics 47: 1059-1069.

Shao, Jun and Shein-Chung Chow. 1994. Statistical inference in stability analysis. Biometrics 50: 753-763.

Tsong, Yi, Wen-Jen Chen and Chi Wan Chen. 2003. ANCOVA approach for shelf life analysis of stability study of multiple factor design. Journal of Biopharmaceutical Statistics 13: 375-393.

BIOGRAPHICAL SKETCH

Youngkyoung Min was born in Seoul, South Korea. She received her bachelor's and master's degree in polymer engineering from Chungnam National University, South Korea. In August 2001, she also achieved her M.S. in the Management Department at the University of Florida. In the fall of 2002, she enrolled for graduate studies in the Statistics Department at the University of Florida and will receive her Master of Science in Statistics in August 2004.