Vol. 2: 60-67

THE UNIVERSITY OF CENTRAL FLORIDA
UNDERGRADUATE RESEARCH JOURNAL

Published
November 2, 2007

Application of Transformations in Parametric Inference

Naomi Brownstein
Faculty Mentor: Dr. Marianna Pensky

ABSTRACT: One can apply transformations of random variables to conduct inference for multiple distributions in a few simple steps. These methods are used routinely in maximum likelihood estimation but are rarely applied in other statistical procedures. In this project, transformations of variables were explored and applied to derivations of the best unbiased estimators, Bayesian estimators, construction of various kinds of priors, estimation and inference in the stress-strength problem. First, general results were obtained on the application of transformations of random variables to the derivation of numerous statistical procedures. Second, common distributions and the relationships between them were listed in a table. Third, examples of applications of our theory were provided; i.e., papers published in various statistical journals were examined and the same results were obtained in just a few lines with almost no effort. The value of this project lies in the fact that undergraduate level statistics can yield such powerful results.

Republication not permitted without written consent of the author.

INTRODUCTION

The method of transformations of random variables has been a standard tool in statistical inference. Transformations have been used for solutions of various statistical problems, such as nonparametric density estimation, nonparametric regression, analysis of time series, and construction of equivariant estimators.[2-7] Journals and books such as "Continuous Univariate Distributions" and "Unbiased Estimators and Their Applications" construct statistical inference for dozens of statistical distributions.[8,9,10] Procedures are usually conducted for each distribution family separately; often, at least one of the numerous calculations required leads to errors.

However, one simple application has been largely overlooked by the statistical community. The objective of this paper is to provide a simple approach to statistical inferences using the method of transformations of variables. It is a well-known fact that the majority of familiar probability distributions are just transformations of one another. Consequently, results of parametric statistical inference for one family of pdfs can be reproduced without much work for another family. We shall demonstrate performance of the powerful tool of transformations on examples of constructions of various estimation procedures, hypothesis testing, Bayes inference, stress-strength problems. We argue that the tool of transformations not only should be used more widely in statistical research but should also become a routine part of calculus-based courses of statistics. The following results include only the case of a one-dimensional random variable. While the theory has an obvious extension to the case of random vectors, making this generalization would unnecessarily complicate the presentation.

Consider a random variable *X* with the pdf $f(x|\ \boldsymbol{\theta})$, where parameter $\boldsymbol{\theta}$ is a scalar or a vector. Suppose also that there exist a random variable ξ , a monotone function u and a one-to-one transformation ν such that $X = u(\xi)$, where the pdf $g(\xi|\ \boldsymbol{\tau})$ of ξ has a different parameterization from *X*, namely,

$$g(\xi|\ \boldsymbol{\tau}) = f(u(\xi)|\nu(\ \boldsymbol{\tau}))\ |u'(\xi)|, \qquad \boldsymbol{\theta} = \nu(\ \boldsymbol{\tau}) \qquad (1)$$

Denoting $u^{-1} = v$ and $\nu^{-1} = \eta$,we rewrite (1) as

$$f(x|\ \boldsymbol{\theta}) = g(v(x)|\eta(\ \boldsymbol{\theta}))\ |v'(x)|, \qquad \boldsymbol{\tau} = \eta(\ \boldsymbol{\theta}). \qquad (2)$$

Now, let $g(\xi|\ \boldsymbol{\tau})$ be a popular distribution family, so that all sorts of statistical results are available. The objective of the present paper is a re-formulation of these results for $f(x|\ \boldsymbol{\theta})$. Here, notice that the correspondence (1) is quite common (see Table 1) but is not used to the full extent. Our goal is not to explore all possible correspondences of this sort but to provide few examples which will illustrate the general idea which can be easily extended to many more kinds of statistical procedures and various other families of distributions. For this reason, this paper focuses on only a few distributions which can be obtained by transformations from a scale or location-scale family of exponential distributions. The majority of those distributions are listed in Table 1.

For example, for the exponential and Weibull distributions, $g(\xi|\tau)$ and , $f(x|\ \boldsymbol{\theta})$ respectively, we have , $v(x) = x^{\alpha}$, $\boldsymbol{\theta} = \sigma$, $\tau = \lambda = \eta(\sigma) = \sigma$; and $g(\xi) = \lambda \exp\{-\lambda\xi\}$ and $\alpha > 0$. Thus

$$\begin{aligned} g(v(x)|\eta(\ \boldsymbol{\theta}))\ |v'(x)| &= g(v(x)|\eta(\sigma))\ |v'(x)| \\ &= \sigma^{-\alpha} \exp\{-\sigma^{-\alpha}(x^{\alpha})\}\ |\alpha x^{\alpha-1}| \\ &= \tfrac{\alpha}{\sigma^{\alpha}} x^{\alpha-1} \exp\left\{-\left(\tfrac{x}{\sigma}\right)^{\alpha}\right\} \\ &= f(x|\sigma) \\ &= f(x|\ \boldsymbol{\theta}) \end{aligned}$$

Voinov and Nikulin list 115 uniform minimum variance unbiased estimators in the case of the one-parameter exponential distribution and only 31 estimators in the case of the Weibull distribution. The use of Table 1 and transformation of variables can yield each of the missing estimators for the Weibull distribution in a few short lines. We will discuss these techniques in the latter part of the present paper.

Some readers of this paper may remark that some of the results listed here can be obtained in a general form for, say, one or two-parameter exponential families. The goal, however, is not to provide such a generalization but to supply a simple and yet powerful methodological tool to modify statistical procedures. The scale and location-scale family of exponential distributions is used here only as an example. In fact, techniques described below can be used for a distribution family that does not have a sufficient statistic.

The rest of the paper is organized as follows. Table 1 showing transformations between distribution families is at the end of this introduction. Section 2 considers basic statistical inference for $f(x|\ \boldsymbol{\theta})$ based on inference for $g(\xi|\tau)$ - sufficient statistics, maximum likelihood estimators (MLE) and uniform minimum variance unbiased estimators (UMVUE), interval estimators and likelihood ratio tests. Section 3 provides examples of results which can be obtained by applying techniques suggested in this paper. Section 4 concludes the paper with the discussion. Section 5 contains proofs of all statements formulated in previous sections.

Table 1: *Transformations of random variables.*

$f(x\|\ \boldsymbol{\theta})$	Tranforms	$g(\xi\|\tau)$
Weibull distribution: $f(x\|\sigma) = \frac{\alpha}{\sigma^\alpha} x^{\alpha-1} \exp\left\{-\left(\frac{x}{\sigma}\right)^\alpha\right\}$, α known, $x > 0$.	$v(x) = x^\alpha$ $\boldsymbol{\theta} = \sigma$ $\boldsymbol{\tau} = \lambda$ $\eta(\sigma) = \sigma^{-\alpha}$	One-parameter exponential distribution: $g(\xi\|\lambda) = \lambda \exp\{-\lambda\xi\}$, $\xi > 0$
Rayleigh distribution: $f(x\|\sigma) = \frac{x}{\sigma} \exp\left\{-\frac{x^2}{2\sigma}\right\}$, $x > 0$.	$v(x) = x^2/2$ $\boldsymbol{\theta} = \sigma$ $\boldsymbol{\tau} = \lambda$ $\eta(\sigma) = \sigma$	One-parameter exponential distribution: $g(\xi\|\lambda) = \lambda \exp\{-\lambda\xi\}$, $\xi > 0$
Burr type X distribution: $f(x\|\sigma) = 2\sigma x e^{-x^2}(1-e^{-x^2})^{\sigma-1}$, $x > 0$.	$v(x) = -\ln(1-e^{-x^2})$ $\boldsymbol{\theta} = \sigma$ $\boldsymbol{\tau} = \lambda$ $\eta(\sigma) = \sigma$	One-parameter exponential distribution: $g(\xi\|\lambda) = \lambda \exp\{-\lambda\xi\}$, $\xi > 0$
Burr type XII distribution: $f(x\|\sigma) = \frac{\alpha x^{\alpha-1}}{\sigma(1+x^\alpha)^{\frac{1+\sigma}{\sigma}}}$, α known, $x > 0$.	$v(x) = \ln(1+x^\alpha)$ $\boldsymbol{\theta} = \sigma$ $\boldsymbol{\tau} = \lambda$ $\eta(\sigma) = \sigma$	One-parameter exponential distribution: $g(\xi\|\lambda) = \lambda \exp\{-\lambda\xi\}$, $\xi > 0$
Extreme Value distribution: $f(x\|\sigma) = \frac{1}{\sigma} \exp\left\{x - \frac{\exp x - 1}{\sigma}\right\}$ $x > 0$.	$v(x) = \exp\{x\} - 1$ $\boldsymbol{\theta} = \sigma$ $\boldsymbol{\tau} = \lambda$ $\eta(\sigma) = \sigma^{-1}$	One-parameter exponential distribution: $g(\xi\|\lambda) = \lambda \exp\{-\lambda\xi\}$, $\xi > 0$

(continued on next page)

Pareto distribution: $f(x\|\sigma,\rho) = \frac{\sigma\rho^{\sigma}}{x^{\sigma+1}} I(x \geq \rho)$.	$v(x) = \ln x$ $\boldsymbol{\theta} = (\sigma,\rho)$ $\boldsymbol{\tau} = (\lambda,\mu)$ $\eta(\sigma,\rho) = (\sigma, \ln\rho)$	Two-parameter exponential distribution: $g(\xi\|\mu,\lambda) = \lambda \exp\{-\lambda(\xi-\mu)\}$, $\xi \geq \mu$.
Power distribution: $f(x\|\sigma,\rho) = \frac{\sigma}{\rho^{\sigma}} x^{\sigma-1} I(0 < x < \rho)$.	$v(x) = \ln(1/x)$ $\boldsymbol{\theta} = (\sigma,\rho)$ $\boldsymbol{\tau} = (\lambda,\mu)$ $\eta(\sigma,\rho) = (\sigma, \ln(1/\rho))$	Two-parameter exponential distribution: $g(\xi\|\mu,\lambda) = \lambda \exp\{-\lambda(\xi-\mu)\}$, $\xi \geq \mu$.

URJ THE UNIVERSITY OF CENTRAL FLORIDA
UNDERGRADUATE RESEARCH JOURNAL
2: 60–67

BASIC STATISTICAL INFERENCE PROCEDURES

In this section construction of the main statistical procedure for $f(x|\ \theta)$ based on the relevant knowledge for $g(\xi|\ \tau)$ is discussed. Let the relation between $f(x|\ \theta)$ and $g(\xi|\ \tau)$ be given by (1) and (2) and $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ and $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$ be the independent and identically distributed (i.i.d.) samples from those pdfs with $X_i = u(\xi_i), i = 1, \cdots, n$. Then the parametric family $f(x|\ \theta)$ inherits all useful properties of the family $g(\xi|\ \tau)$. To simplify the notations in what follows $v(\mathbf{X}) = (v(X_1), v(X_2), \cdots, v(X_n))$.

Theorem 1 Let $T(\ \xi) = T(\xi_1, \xi_2, \cdots, \xi_n)$ be a scalar or vector valued sufficient statistic for the family of pdfs $g(\xi|\ \tau)$. Then $T^*(\mathbf{X}) = T(v(\mathbf{X}))$ is a sufficient statistic for the family of pdfs $f(x|\ \theta)$.

Theorem 2 Let $U(\ \xi) = U(\xi_1, \xi_2, \cdots, \xi_n)$ be a MLE of τ based on the sample ξ. Then $U^*(\mathbf{X}) = \nu(U(v(\mathbf{X})))$ is the MLE of θ based on observations $\mathbf{X}$. If, moreover, $T = T(\ \xi)$ is a sufficient statistic τ for , so that the MLE of τ has the form $U(\ \xi) = V(T)$, then the MLE of θ is where $T^* = T^*(\mathbf{X})$ is defined in Theorem 1.

Theorem 3 Let $h(\ \tau)$ be a function of τ and let $V(T)$ be the UMVUE of $h(\ \tau)$ based on the sufficient statistic $T = T(\ \xi)$. Then $V(T^*)$ is the UMVUE of $h^*(\ \theta) = h(\eta(\ \theta))$ based on the sufficient statistic $T^*(\mathbf{X})$.

Corollary 1 Let $\Psi(\xi_1, \cdots, \xi_n)$ be a function of observations $\xi_1, \cdots, \xi_n$ with expectation $\mathrm{E}_{\ \tau}\Psi(\xi_1, \cdots, \xi_n)$ over the pdf $\prod_{j=1}^n g(\xi_j|\ \tau)$. If $V(T)$ is the UMVUE of $\mathrm{E}_{\ \tau}\Psi(\xi_1, \cdots, \xi_n)$, then $V(T^*)$ is the UMVUE of $\mathrm{E}_{\ \theta}\Psi(v(X_1), \cdots, v(X_n))$.

Theorems 1-3 and Corollary 1 refer to construction of various types of point estimators. However, interval estimation and hypothesis testing procedures can be modified in a similar way.

Theorem 4 Let $\varphi(\ \tau)$ be a parametric function of interest and let $(\underline{U}(\ \xi), \overline{U}(\ \xi))$ be an interval estimator of $\varphi(\ \tau)$ corresponding to the confidence level $(1-\gamma)$, i.e.

$$P(\underline{U}(\ \xi) < \varphi(\ \tau) < \overline{U}(\ \xi)) \geq 1 - \gamma.$$

Let $\underline{U}^*(\mathbf{X}) = \underline{U}(v(\mathbf{X}))$ and $\overline{U}^*(\mathbf{X}) = \overline{U}(v(\mathbf{X}))$. Then $(\underline{U}^*(\mathbf{X}), \overline{U}^*(\mathbf{X}))$ is an interval estimator of $\varphi(\eta(\ \theta))$ corresponding to the confidence level $(1-\gamma)$.

Theorem 5 Let $U(\ \xi) = U(\xi_1, \xi_2, \cdots, \xi_n)$ be a likelihood ratio test (LRT) statistic for testing hypothesis $H_0:\ \tau \in \Omega_0$ versus $H_1:\ \tau \in \Omega_0^C$. Consider sets Θ_0 and Θ_0^C such that

$$\tau \in \Omega_0 \Leftrightarrow \theta = \nu(\ \tau) \in \Theta_0, \qquad \tau \in \Omega_0^C \Leftrightarrow \theta = \nu(\ \tau) \in \Theta_0^C$$

Then $U^*(\mathbf{X}) = U(v(\mathbf{X}))$ is a LRT statistic for testing hypothesis $H_0^*:\ \theta \in \Theta_0$ versus $H_1^*:\ \theta \in \Theta_0^C$. Moreover, if $U(\ \xi) > C_\gamma$ is a level γ test for H_0, then $U^*(\mathbf{X}) > C_\gamma$ is a level γ test for H_0^*.

EXAMPLES

This section provides several examples of applications of the theory presented.

Example 1. Consider the task of finding the UMVUE $\ln(\sigma)$ of given a sample $\mathbf{X}$ from the Weibull distribution in Table 1. Using the fact that $\ln(\sigma) = -1/\alpha \ln(\lambda)$ yields $\sigma = \nu(\lambda) = \lambda^{-1/\alpha}$. The UMVUE of $\ln(\lambda)$ based on the sample ξ from the one-parameter exponential distribution has the same form as in [9].

$$\mathrm{V(T)} = \ln(T) - \Psi(n) \qquad (3)$$

where $T = \bar{\xi}$ and $\Psi(n) =$ is the Euler's psi-function $\Psi(x) = \frac{d}{dx}\ln\Gamma(x)$. Theorems 1 and 3 with $v(x) = x^\alpha$ imply that the UMVUE for $\ln(\sigma)$ is $V(T^*) = -1/\alpha[\ln(T) - \Psi(n)]$ where $T^* = n^{-1}\sum_{i=1}^n X_i^\alpha$.

Note that the formula for UMVUE of $\ln(\sigma)$ is not listed in the table of UMVUEs for the Weibull distribution nor are estimators for many other functions of σ which can be easily obtained by using our simple technique.[10] Moreover, the method of transformations of random variables is not even listed among half a dozen techniques suggested for derivation of UMVUEs. If the author introduced this very simple idea in the book the Tables would be more comprehensive, while the book would be much shorter.

Example 2. The tool of transformations is very useful for solutions to simple problems in various statistics courses. For example, consider problem 8.5 which we formulate using the same notation as in the present paper.[11] The objective is to construct the MLE for parameters (σ, ρ) of a Pareto population as well as the LRT for $H_0 : \rho = 1$ versus $H_1 : \rho \neq 1$ in the presence of an unknown parameter ρ.
Using the fact that the MLE for (λ, μ) in the case of the two-parameter exponential distribution is $(1/ \bar{\boldsymbol{\xi}}, \xi_{(1)}$ and the relations $v(x) = \ln x$ and $\nu(\mu) = e^{\mu}$, the author derives from Theorem 2 that the MLE for (σ, ρ) is of the form

$$\hat{\sigma} = n \left(\sum_{i=1}^{n} \ln(X_i) \right)^{-1}, \quad \hat{\rho} = \exp(\ln(X_{(1)})) = X_{(1)}.$$

Because the relationship between μ and ρ testing H_0 is equivalent to $H_0^* : \mu = 0$. The LRT test for H_0^* has the acceptance region $c_1 < U(\boldsymbol{\xi}) < c_2$ where $U(\boldsymbol{\xi}) = \bar{\boldsymbol{\xi}}/\xi_{(1)}$, so that the LRT test for H_0 has the acceptance region $c_1 < U^*(\mathbf{X}) < c_2$ where

$$U^*(\mathbf{X}) = n^{-1} \sum_{i=1}^{n} \ln(X_i)/\ln(X_{(1)}).$$

Also, the distributions of $U(\boldsymbol{\xi})$ and $U^*(\mathbf{X})$ are the same.

DISSCUSSION

In the present paper, the tool of transformations of random variables is exploited to obtain various results in statistical inference with minimal effort. We have discussed examples of straightforward applications of the transformations, all of which can be appreciated by a student who has taken an upper level undergraduate or lower level graduate statistics course. Nevertheless, an example concerning unbiased estimation would allow one to expand the list of UMVUEs significantly with minimal effort.

However, the utility of the method is not just limited to just the few cases listed in Section 2. It can be applied to finding Bayes and empirical Bayes estimators, constructing Bayesian credible sets, estimation and testing in the stress-strength model 1 and so on. It can be argued that since the method of transformations is a routine part of almost any calculus-based statistics course, it would be worth introducing the methodology discussed above into standard textbooks, thereby saving not only hundreds of man-hours but also hundreds of trees.

PROOFS

Proof of Theorem 1. Let $T(\boldsymbol{\xi})$ be a sufficient statistic for the family of pdfs $g(\xi|\tau)$. Then, by the factorization theorem, $\prod_{i=1}^{n} g(\xi_i|\theta) = w(\boldsymbol{\xi}) q(T(\boldsymbol{\xi}), \boldsymbol{\tau})$ for some functions $w(\cdot)$ and $q(\cdot,\cdot)$. The latter implies

$$\begin{aligned} \prod_{i=1}^{n} f(X_i|\theta) &= \prod_{i=1}^{n} \left[g(v(X_i)|\eta(\boldsymbol{\theta})) |v'(X_i)| \right] \\ &= q(T(v(\mathbf{X})), \eta(\boldsymbol{\theta})) w(v(\mathbf{X})) \prod_{i=1}^{n} |v'(X_i)| \\ &= q(T(v(\mathbf{X})), \eta(\boldsymbol{\theta})) w^*(v(\mathbf{X})) \\ &= q(T^*((\mathbf{X}), \eta(\boldsymbol{\theta})) w^*(v(\mathbf{X})), \end{aligned}$$

so that is $T^*((\mathbf{X}))$ the sufficient statistic for $\boldsymbol{\theta}$.

Proof of Theorem 2. Let $L(\boldsymbol{\tau}| \boldsymbol{\xi})$ be the likelihood for $\boldsymbol{\tau}$, $\hat{\boldsymbol{\tau}} = \arg\max_{\boldsymbol{\tau}} L(\boldsymbol{\tau}| \boldsymbol{\xi})$, and define μ such that $\mu(f(\hat{\alpha})) = \arg\max_{\alpha} L(f(\alpha)|v(\mathbf{X}))$ for all $f(\hat{\alpha})$. Recall that $\hat{\boldsymbol{\tau}} = \eta(\hat{\boldsymbol{\theta}})$. Then the likelihood for $\boldsymbol{\theta}$ is of the form

$$L^*(\boldsymbol{\theta}|\mathbf{X}) = \prod_{i=1}^{n} g(v(X_i)|\eta(\boldsymbol{\theta})) = L(\eta(\boldsymbol{\theta})|v(\mathbf{X})),$$

so that

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} L(\eta(\boldsymbol{\theta})|v(\mathbf{X})) = \mu(\eta(\hat{\boldsymbol{\theta}})) = \mu(\hat{\boldsymbol{\tau}}).$$

To complete the proof, use the invariance property of the MLE and Theorem whenever there exists a sufficient statistic for $\boldsymbol{\tau}$.

Proof of Theorem 3. Let $T(\boldsymbol{\xi})$ be a sufficient statistic for $\boldsymbol{\tau}$ and let $V(T)$ be a UMVUE of $h(\boldsymbol{\tau})$. Hence

$$\begin{aligned} h(\eta(\boldsymbol{\theta})) &= h(\boldsymbol{\tau}) = \int \cdots \int V(T(\boldsymbol{\xi})) \prod_{i=1}^{n} g(\xi_i| \boldsymbol{\tau}) d \boldsymbol{\xi} \\ &= \int \cdots \int V(T(v(\mathbf{X})) \prod_{i=1}^{n} f(v(X_i)|\eta(\boldsymbol{\theta})) d\mathbf{X} \end{aligned}$$

and $V(T(v(\mathbf{X})) = V(T^*(\mathbf{X}))$ is the UMVUE of $h(\eta(\boldsymbol{\theta}))$.

Proof of Theorem 4: Let $(\underline{U}(\boldsymbol{\xi}), \overline{U}(\boldsymbol{\xi}))$ be an interval estimator of $\varphi(\boldsymbol{\tau})$ corresponding to the confidence level $(1-\gamma)$, i.e.
$P(\underline{U}(\boldsymbol{\xi}) < \varphi(\boldsymbol{\tau}) < \overline{U}(\boldsymbol{\xi})) \geq 1 - \gamma.$ Hence,

$$1-\gamma \;\leq\; \int\int \prod_{i=1}^{n} g(\xi_i|\tau) I(\underline{U}(\boldsymbol{\xi}) \leq \nu(\tau) \leq \overline{U}(X)) dx_1 \ldots dx_n$$

$$= \int\int \prod_{i=1}^{n} g(v(x_i)|\tau)|v'(x_i)| I(\underline{U}(v(x_1),\ldots,v(x_n))$$

$$\leq \varphi(\nu(\tau) \leq \overline{U}(v(x_1),\ldots,v(x_n)) dx_1 \ldots dx_n$$

$$= \int\int \prod_{i=1}^{n} f(x_i|\theta) I(\underline{U}^*(xi) \leq \varphi(\nu(\theta) \leq \overline{U}^*(X) dX$$

Therefore, $[\underline{U}^*(X), \overline{U}^*(X)]$ is the confidence interval for $\varphi(\nu(\theta))$ with the confidence level $1-\gamma$.

Proof of Theorem 5: Let $U(\boldsymbol{\xi}) = U(\xi_1, \xi_2, \cdots, \xi_n)$ be a LRT statistic testing $H_0: \boldsymbol{\tau} \in \Omega_0$ versus , $H_1: \boldsymbol{\tau} \in \Omega_0^C$ and $\boldsymbol{\tau} \in \Omega_0 \Leftrightarrow \boldsymbol{\theta} = \nu(\boldsymbol{\tau}) \in \Theta_0$, $\boldsymbol{\tau} \in \Omega_0^C \Leftrightarrow \boldsymbol{\theta} = \nu(\boldsymbol{\tau}) \in \Theta_0^C$. Then,

$$U(\boldsymbol{\xi}) = \frac{sup_{\tau\in\Omega_0} \prod_{i=1}^{n} g(\xi_i|\tau)}{sup_{\tau\in\Omega} \prod_{i=1}^{n} g(\xi_i|\tau)}$$

$$= \frac{sup_{\theta\in\Theta_0} \prod_{i=1}^{n} g(\xi_i|\eta(\theta))}{sup_{\theta\in\Theta} \prod_{i=1}^{n} g(\xi_i|\eta(\theta))}$$

$$= \frac{sup_{\theta\in\Theta_0} \prod_{i=1}^{n} g(v(x_i)|\eta(\theta)) \prod_{i=1}^{n} |v'(x_i)|}{sup_{\theta\in\Theta} \prod_{i=1}^{n} g(v(x_i)|\eta(\theta)) \prod_{i=1}^{n} |v'(x_i)|}$$

$$= \frac{sup_{\theta\in\Theta_0} \prod_{i=1}^{n} f(x_i|\theta)}{sup_{\theta\in\Theta} \prod_{i=1}^{n} f(x_i|\theta)} = U(v(X))$$

Hence $U(v(X))$ is the desired LRT statistic for H_0^*.

2: 60–67

REFERENCES

1. N. Brownstein and M. Pensky. "Estimation in the Stress-Strength Model."

2. S. Konishi, "Normalizing Transformations and Bootstrap Confidence Intervals," *Ann. Statist.* 19, (1991): 2209-2225.

3. O. B. Linton, R. Chen, N. Wang, and W. Härdle, "An Analysis of Transformations for Additive Nonparametric Regression," *J. Amer. Statist. Assoc.* 92, (1997): 1512-1521.

4. J. S. Marron, and D. Ruppert, "Transformations to Reduce Boundary Bias in Kernel Density Estimation." *J. Roy. Statist. Soc. Ser.* B 56, (1994): 653-671.

5. D. Ruppert, and D. B. H. Cline, "Bias Reduction in Kernel Density Estimation by Smoothed Empirical Transformations." *Ann. Statist. 22*, (1994): 185-210.

6. S. Van Buuren, "Optimal Transformations for Categorical Autoregressive Time Series." *Statist. Neerlandica 51*, (1997): 90-106.

7. L. Yang, J. S. Marron, "Iterated Transformation-kernel Density Estimation." *J. Amer. Statist. Assoc.* 94, (1999): 580-589.

8. N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions.*, Vol. 1. 2nd ed. (Wiley, New York: 1994).

9. N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions.*, Vol. 2. 2nd ed. (Wiley, New York: 1995).

10. V. G. Voinov and M. S. Nikulin, *Unbiased Estimators and Their Applications.* Vol. 1: Univariate Case. (Dordrecht, Netherland: Kluwer Academic Publishers, 1993).

11. G. Casella, R. L. Berger, *Statistical Inference.* 2nd ed. (Duxbury, California: 2002).