Journal of the Korean
Statistical Society
Vol. 26, No. 3, 1997

On the Conditional Tolerance Probability in Time Series Models †

Sangyeol Lee [1]

Abstract

Suppose that $\{X_i\}$ is a stationary AR(1) process and $\{Y_j\}$ is an ARX process with $\{X_i\}$ as exogeneous variables. Let Y_j^* be the stochastic process which is the sum of Y_j and a nonstochastic trend. In this paper we consider the problem of estimating the conditional probability that Y_{n+1}^* is bigger than X_{n+1}, given $X_1, Y_1^*, \ldots, X_n, Y_n^*$. As an estimator for the tolerance probability, an Mann-Whitney statistic based on least squares residuals is suggested. It is shown that the deviations between the estimator and true probability are stochastically bounded with $n^{-1/2}$ order. The result may be applied to the stress-strength reliability theory when the stress and strength variables violate the classical iid assumption.

Key Words : Tolernace probability; Mann-Whitney statistic; AR(1) stationary process; ARX process; Stress-strength reliability theory; Berstein's inequality for martingales.

†This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1996.

[1]Department of Statistics, Seoul National University, Seoul, 151-742, Korea.

1. INTRODUCTION

Let $\{X_i\}$ be the stationary AR(1) process of the form

$$X_i - \mu = \alpha(X_{i-1} - \mu) + \delta_i, \quad |\alpha| < 1, \tag{1.1}$$

where δ_i are iid $(0, \sigma^2)$ from a continuous distribution F, and $\{Y_j\}$ be the stationary ARX process, such that

$$Y_j = \beta Y_{j-1} + \gamma(X_{j-1} - \mu) + \varepsilon_j, \quad |\beta| < 1, \tag{1.2}$$

where ε_j are iid $(0, \tau^2)$ from a continuous distribution G and independent of $\{\delta_i\}$. Let $Y_j^* = \omega_j + Y_j$, where ω_j is a nonstochastic trend. The ARX models are widely used in analyzing linear systems: see Hannan and Diestler (1988) for the definition and applications of ARX models.

The issue of this paper is motivated by stress-strength reliability theory despite the technical result itself (cf. Theorem 2.1) is of our interest. In the classical stress-strength reliability theory both the stress X and the strength Y are treated as random variables, and there is a concern for estimating the reliability $P(X < Y)$. See, for example, Birnbaum (1956), Govindarajulu (1968), Helperin et. al. (1987) and Guttman et. al. (1988). Usually, in nonparametric setting, the reliability is estimated through the Mann-Whitney statistic assuming that the observations X and Y's are mutually independent and iid. However, in real practice, it is possible that both the stress and strength random variables are serially correlated and the strength is severely affected by the stress. For instance, one can consider the situation where the stress and strength random variables are associated with some part of a system and they follow the models in (1.1) and (1.2). Suppose that the system is well-functioning up to time n and that one would like to find a rule, based on current observations $X_i, Y_i^*, i = 1, \ldots, n$, for deciding whether the part under consideration should be replaced or not. In this case, it is natural to adopt the decision rule on which the part is replaced if $p_n = P(X_{n+1} < Y_{n+1}^* | X_i, Y_i^*, i = 1, \ldots, n)$ is less than a preassigned number $p_0 \in (0, 1)$. However, since p_n is unknown, the decision rule should be based on an estimate of p_n.

Assume that $X_0 = 0$ and $Y_0 = 0$. Further, for simplicity, assume that the trend ω_j is known. Suppose that given observations $X_1, Y_1^*, \ldots, X_n, Y_n^*$, one wishes to estimate the conditional probability of the event that Y_{n+1}^* is bigger than X_{n+1}. Let

$$p_n = P(X_{n+1} < Y_{n+1}^* | X_i, Y_i^*, i = 1, \ldots, n), \tag{1.3}$$

and set $\underset{\sim}{\theta} = (\mu, \alpha, \beta, \gamma)'$. Due to the equations in (1.1) and (1.2), we can write

$$\begin{aligned} p_n &= P(X_{n+1} < Y^*_{n+1} | X_n, Y_n) \\ &= H(\xi_n), \end{aligned}$$

where H is the common distribution of $\delta_i - \varepsilon_j$ and $\xi_n = \beta Y_n + (\gamma - \alpha)X_n - \mu + \omega_{n+1}$. Therefore, provided that δ_i and ε_j are observable and $\underset{\sim}{\theta}$ is known, p_n can be estimated by the Mann-Whitney statistic

$$n^{-2} \sum_{i,j=1}^{n} I(\delta_i - \varepsilon_j < \xi_n). \tag{1.4}$$

Since the errors are unobservable and the parameters are unknown, we estimate p_n based on residuals.

Let $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ and let $\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n$ be the estimators of α, β, γ based on X_i, Y_i, $i = 1, \ldots, n$, such that $\hat{\alpha}_n - \alpha = O_P(n^{-1/2})$, $\hat{\beta}_n - \beta = O_P(n^{-1/2})$ and $\hat{\gamma}_n - \gamma = O_P(n^{-1/2})$. Define the residuals $\hat{\delta}_i = X_i - \bar{X} - \hat{\alpha}_n(X_{i-1} - \bar{X})$ and $\hat{\varepsilon}_j = Y_j - \hat{\beta}_n Y_{j-1} - \hat{\gamma}_n(X_{j-1} - \bar{X})$. Then, the Mann-Whitney statistic based on the residuals is $\hat{p}_n(\hat{\xi}_n)$, where

$$\hat{p}_n(x) = n^{-2} \sum_{i,j=1}^{n} I(\hat{\delta}_i - \hat{\varepsilon}_j < x), \quad x \in R, \tag{1.5}$$

and

$$\hat{\xi}_n = \hat{\beta}_n Y_n + (\hat{\gamma}_n - \hat{\alpha}_n)X_n - \bar{X} + \omega_{n+1}. \tag{1.6}$$

In Section 2, it is shown that under regularity conditions,

$$p_n - \hat{p}_n(\hat{\xi}_n) = O_P(n^{-1/2})$$

(cf. Theorem 2.1). For proving the above, the following turns out to be very useful:

$$\sup_x |\hat{p}_n(x) - H(x)| = O_P(n^{-1/2}).$$

In fact, the above is much related to the oscillation result on the residual empirical processes in stationary autoregressive processes (cf. Boldin (1982)). Like Lee and Wei (1996), Bernstein's inequaluty for martingales plays an importnat role.

Although we do not present details here, it can be shown without difficulties that Theorem 2.1 is extended to high dimensional cases where $\{X_i\}$ is the stationary AR(p) process and $\{Y_j\}$ is the stationary ARX process, such that

$$Y_j = \beta_1 Y_{j-1} + \cdots + \beta_q Y_{j-q} + \gamma_1 (X_{j-1} - \mu) + \ldots + \gamma_r (X_{j-r} - \mu) + \varepsilon_j.$$

The proof in this case is essentially the same as of Theorem 2.1.

2. MAIN RESULTS

Theorem 2.1. Let p_n, $p_n(x)$ and $\hat{\xi}_n$ be the same as in (1.3), (1.5) and (1.6). If F and G satisfy

$$\sup_x \{|F'(x)| + |F''(x)| + |G'(x)| + |G''(x)|\} < \infty, \tag{2.1}$$

then

$$|p_n - \hat{p}_n(\hat{\xi}_n)| = O_P(n^{-1/2}). \tag{2.2}$$

Before we prove Theorem 2.1, we introduce lemmas.

Lemma 2.1. Assume that (2.1) holds. Then for any $K > 0$,

$$\sup_{x,|a|\le K} |n^{-3/2} \sum_{i,j=1}^{n} \{F(x + \varepsilon_j + an^{-1/2}(X_{i-1} - \mu)) \\ - H(x + an^{-1/2}(X_{i-1} - \mu)) + H(x) - F(\varepsilon_j + x)\}| = o_P(1).$$

Proof. By Taylor's series expansion, one can show that

$$\begin{aligned} &\sup_{x,|a|\le K} |n^{-3/2} \sum_{i,j=1}^{n} \{F(x + \varepsilon_j + an^{-1/2}(X_{i-1} - \mu)) - F(x + \varepsilon_j)\}| \\ \le\ & A\{|n^{-1} \sum_{i=1}^{n} (X_i - \mu)| + n^{-3/2} \sum_{i=1}^{n} (X_{i-1} - \mu)^2\}, \quad A > 0, \end{aligned}$$

which goes to 0 in probability. Similarly, it can be shown that

$$\sup_{x,|a|\le K} |n^{-3/2} \sum_{i,j=1}^{n} \{H(x + an^{-1/2}(X_{i-1} - \mu)) - H(x)\}| = o_P(1).$$

Therefore, the lemma is established. □

Lemma 2.2. Assume that F and G satisfy (2.1). Put

$$\begin{aligned}\Gamma_n(x,b) &= \sup_{x,|b|\le K} n^{-1/2}\Big|\sum_{j=1}^{n}\{F(x+\varepsilon_j+bn^{-1/2}Y_{j-1})-H(x+bn^{-1/2}Y_{j-1}) \\ &\quad +H(x)-F(x+\varepsilon_j)\}\Big| = o_P(1). \end{aligned} \tag{2.3}$$

Then for any $K>0$, $\sup_{x,|b|\le K}|\Gamma_n(x,b)| = o_P(1)$.

The following lemma, due to Shorack and Wellner (1986, P. 809), turns out to be useful to prove Lemma 2.2.

Lemma 2.3. (Bernstein's inequality for martingale differences). Let $\{\sum_{i=1}^{n} X_i, \mathcal{F}_n; n\ge 1\}$ be a martingale with $EX_1 = 0$ and τ is a stopping time. Suppose that there exists positive constants M and V such that $|X_i|\le M$ a.s. for all i and $\sum_{i=1}^{\tau} E(X_i^2|\mathcal{F}_{i-1})\le V$ a.s. Then for any $\eta>0$,

$$P(|\sum_{i=1}^{\tau} X_i|\ge\eta)\le 2\exp\{-\eta^2/2(V+M\eta/3)\}.$$

Proof of Lemma 2.2. Put $W_j = n^{-1/2}Y_j$. Let $N_n = n^2$ and partition the real line by the points

$$-\infty = x_{n0} < x_{n1} < \cdots < x_{nN_n} = \infty$$

such that $H(x_{ni}) = i/N_n, i = 0,1,\ldots,N_n$. For each $x\in(x_{nr},x_{n,r+1}]$, we have

$$|H(x)-H(x_{ni})|\le N_n^{-1},\quad i=r,r+1. \tag{2.4}$$

Then, we can write $|\Gamma_n(x,b)|\le\Gamma_{n1}(x,b)+\Gamma_{n2}(x,b)+\Gamma_{n3}(x,b)$, where

$$\begin{aligned}\Gamma_{n1}(x,b) &= \sup_{i=r,r+1}\Big|n^{-1/2}\sum_{j=1}^{n}[F(\varepsilon_j+x_{ni}+bW_{j-1}) \\ &\qquad -H(x_{ni}+bW_{j-1})+H(x_{ni})-F(\varepsilon_j+x_{ni})]\Big| \\ \Gamma_{n2}(x,b) &= \sup_{i=r,r+1}\Big|n^{-1/2}\sum_{j=1}^{n}[H(x_{ni}+bW_{j-1})-H(x+bW_{j-1})]\Big| \\ \Gamma_{n3}(x,b) &= \sup_{i=r,r+1}\Big|n^{-1/2}\sum_{j=1}^{n}[F(\varepsilon_j+x_{ni})-H(x_{ni})+H(x)-F(\varepsilon_j+x)]\Big|.\end{aligned}$$

By the mean value theorem, (2.1) and (2.4), we have that $\sup_{x,|b|\le K}\Gamma_{n3}(x,b) = o_P(1)$. To deal with $\Gamma_{n2}(x,b)$, use Taylor's series expansion to get

$$H(x_{ni}+bW_{j-1}) = H(x_{ni})+bW_{j-1}H'(x_{ni})+2^{-1}b^2W_{j-1}^2H''(\zeta_j)$$

and

$$H(x+bW_{j-1}) = H(x) + bW_{j-1}H'(x) + 2^{-1}b^2W_{j-1}^2H''(\tilde{\zeta}_j),$$

where ζ_j and $\tilde{\zeta}_j$ are numbers between x_{ni}, $x_{ni}+bW_{j-1}$ and x, $x+bW_{j-1}$, respectively. By simple algebras, we can see that $\sup_{x,|b|\le K}\Gamma_{n2}(x,b) = o_P(1)$.

Now, it remains to deal with $\Gamma_{n1}(x,b)$. Define

$$S_n = (\sum_{j=1}^{n}|W_{j-1}| \le n^{1/2}L,\ \max_{1\le j\le n}|W_{j-1}| \le L).$$

For any $\eta > 0$, we can choose $L > 0$ such that $P(S_n^c) < \eta$ for all n. Note that for $\lambda > 0$,

$$\begin{aligned} P(\sup_{x,|b|\le K}|\Gamma_{n1}(x,b)| > \lambda) &\le P(\sup_{x,|b|\le K}|\Gamma_{n1}(x,b)| > \lambda, S_n) + \eta \\ &\le P\left(\sup_{\substack{0\le r\le N_n \\ |b|\le K}}|n^{-1/2}\sum_{j=1}^{n}d_j(x_{nr},b)| > \lambda, S_n\right) + \eta, \end{aligned}$$

where

$$d_j(x,b) = F(\varepsilon_j + x + bW_{j-1}) - H(x+bW_{j-1}) + H(x) - F(\varepsilon_j + x).$$

Hence, it suffices to show that

$$\sup_{\substack{0\le r\le N_n \\ |b|\le K}}\left|n^{-1/2}\sum_{j=1}^{n}d_j(x_{nr},b)\right|I(S_n) = o_P(1). \tag{2.5}$$

For this prupose, partition the interval $[-K,K]$ by the points $b_s = -K + 2Ks/n^2$, $s = 0,\ldots,n^2$, and let $w_{js}^+ = \sup_{b_s\le b<b_{s+1}} bW_j$ and $w_{js}^- = \inf_{b_s\le b<b_{s+1}} bW_j$. For each $b \in [b_s, b_{s+1})$,

$$\begin{aligned} d_j(x,y) &\le F(\varepsilon_j + x + w_{j-1,s}^+) - H(x + w_{j-1,s}^+) + H(x) - F(\varepsilon_j + x) \\ &+ H(x + w_{j-1,s}^+) - H(x + bW_{j-1}), \end{aligned}$$

and similarly,

$$\begin{aligned} d_j(x,y) &\ge F(\varepsilon_j + x + w_{j-1,s}^-) - H(x + w_{j-1,s}^-) + F(x) - H(\varepsilon_j + x) \\ &+ H(x + w_{j-1,s}^-) - F(x + bW_{j-1}). \end{aligned}$$

Since on S_n,

$$\begin{aligned}|H(x+w^{+}_{j-1,s})-H(x+bW_{j-1})| &\le \sup_x |H'(x)|(b_{s+1}-b_s)\max_{1\le j\le n}|Z_{j-1}| \\ &\le n^{-2}M \quad \text{for some } M>0,\end{aligned}$$

and the above also holds for $H(x+w^{-}_{j-1,s})-H(x+yW_{j-1})$, the arguments in (2.5) is bounded by $\Gamma^{*}_{n1}+\Gamma^{*}_{n2}+n^{-3/2}M$, where

$$\Gamma^{*}_{n1}=\sup_{\substack{0\le r\le N_n\\ 0\le s\le n^2}}\Big|n^{-1/2}\sum_{j=1}^{n}d_j(x_{nr},w^{+}_{j-1,s})\Big|I(S_n)$$

and

$$\Gamma^{*}_{n2}=\sup_{\substack{0\le r\le N_n\\ 0\le s\le n^2}}\Big|n^{-1/2}\sum_{j=1}^{n}d_j(x_{nr},w^{-}_{j-1,s})\Big|I(S_n).$$

Here, we only deal with Γ^{*}_{n1} since the other case can be handled in a similar way. Note that w^{+}_{js} is $\mathcal{F}_j:=\sigma(\delta_i,\varepsilon_i,i\le j)$-measurable. Let

$$e_j:=e_j(x_{nr},w^{+}_{j-1,s}):=d_j(x_{nr},w^{+}_{j-1,s})I(\sum_{i=1}^{j}|W_{i-1}|\le n^{1/2}L).$$

Then $\{e_j,\mathcal{F}_j;j\ge 1\}$ is a sequence of martingale differences with $|e_j|\le 1$ and $\sum_{j=1}^{n}E(e_j^2|\mathcal{F}_{j-1})\le n^{1/2}\theta$ for some $\theta>0$, since

$$\begin{aligned}\sum_{j=1}^{n}E(e_j^2|\mathcal{F}_{j-1}) &\le \sum_{j=1}^{n}E(d_j^2|\mathcal{F}_{j-1}) \\ &\le |H(x_{nr}+w^{+}_{j-1,s})-H(x_{nr})|I(\sum_{i=1}^{j}|W_{i-1}|\le n^{1/2}L) \\ &\le \sup_x|H'(x)||w^{+}_{j-1,s}|I(\sum_{i=1}^{j}|W_{i-1}|\le n^{1/2}L) \\ &\le n^{1/2}KL\sup_x|H'(x)|.\end{aligned}$$

Due to Lemma 2.3, we have for all $\lambda>0$,

$$P(|\sum_{j=1}^{n}e_j|>n^{1/2}\lambda)\le 2\exp\{-n\lambda^2/(n^{1/2}\theta+n^{1/2}\lambda/3)\}.$$

Since

$$\begin{aligned} & P(e_j \neq d_j(x_{nr}, w^+_{j-1,s}) \text{ for some } j \leq n \text{ on } S_n) \\ \leq \quad & P(\sum_{j=1}^n |W_{j-1}| > n^{1/2}L \text{ on } S_n) = 0, \end{aligned}$$

we have for all $\lambda > 0$,

$$P(\Gamma^*_{n1} \geq \lambda) \leq 2(n^2+1)(N_n+1)\exp\{-n^{1/2}\lambda^2/(\theta+\lambda/3)\} \to 0,$$

which asserts (2.5). This completes the proof. □

The following lemma can be proved by the arguments of Billingsley (1968, P. 106-107).

Lemma 2.4. Suppose that $\{\eta_n\}$ is a sequence of positive random variables decaying to 0 in probability. Then, provided that $\sup_x\{|F'(x)|+|F''(x)|\} < \infty$ holds, we have

$$\sup_{|x-y|\leq\eta_n} |n^{-1/2}\sum_{i=1}^n\{I(\delta_i < x) - F(x) + F(y) - I(\delta_i < y)\} = o_P(1).$$

The following lemma can be found in Boldin (1982).

Lemma 2.5. Assume that $\sup_x\{|F'(x)| + |F''(x)|\} < \infty$. Then,

$$\begin{aligned} & \sup_x |n^{-1/2}\sum_{i=1}^n\{I(\delta_i < x + bn^{-1/2}X_{i-1}) \\ - \quad & F(x+bn^{-1/2}X_{i-1}) + F(x) - I(\delta_i < x)\}| = o_P(1). \end{aligned}$$

Proof of Theorem 2.1. Observe that

$$\begin{aligned} I(\hat{\delta}_i - \hat{\varepsilon}_j < x) \quad = \quad & I(\delta_i - \varepsilon_j < x + (\hat{\alpha}_n - \alpha + \gamma - \hat{\gamma}_n)(X_{i-1} - \mu) \\ & +(\beta - \hat{\beta}_n)Y_{j-1} + (\bar{X} - \mu)(1 - \hat{\alpha}_n - \hat{\gamma}_n)). \end{aligned} \tag{2.6}$$

Put $Z_j = n^{-1/2}(X_j - \mu)$ and $W_j = n^{-1/2}Y_j$. Since $\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n$ are $n^{1/2}$-consistent estimators, in view of (2.6), we only need to show that for any $K > 0$,

$$\sup_{x,|a|\leq K,|b|\leq K} |n^{-3/2}\sum_{i,j=1}^{n^2} A_{ijk}(x,a,b)| = o_P(1), \quad k = 1,2,3,4,$$

where

$$\begin{aligned}
A_{ij1}(x,a,b) &= I(\delta_i < x+\varepsilon_j + aZ_{i-1} + bW_{j-1}) - F(x+\varepsilon_j + aZ_{i-1} + bW_{j-1}) \\
&\quad + F(x+\varepsilon_j + bW_{j-1}) - I(\delta_i < x+\varepsilon_j + bW_{j-1}) \\
A_{ij2}(x,a,b) &= I(\delta_i < x+\varepsilon_j + bW_{j-1}) - F(x+\varepsilon_j bW_{j-1}) \\
&\quad + F(x+\varepsilon_j) - I(\delta_i < x+\varepsilon_j) \\
A_{ij3}(x,a,b) &= F(x+\varepsilon_j + aZ_{i-1} + bW_{j-1}) - H(x + aZ_{i-1} + bW_{j-1}) \\
&\quad + H(x + aZ_{i-1}) - F(x+\varepsilon_j + aZ_{i-1}) \\
A_{ij4}(x,a,b) &= F(x+\varepsilon_j + aZ_{i-1}) - H(x+aZ_{i-1}) + H(x) - F(x+\varepsilon_j).
\end{aligned}$$

Note that

$$\begin{aligned}
&\sup_{x,|a|\le K,|b|\le K} |n^{-3/2}\sum_{i,j=1}^{n^2} A_{ij1}(x,a,b)| \\
&\le \sup_x |n^{-1/2}\sum_{j=1}^{n}\{I(\delta_i < x + aZ_{i-1}) - F(x+aZ_{i-1}) + F(x) - I(\delta_i < x)\}|,
\end{aligned}$$

which goes to 0 due to Lemma 2.5. On the other hand, we have that

$$\begin{aligned}
&\sup_{x,|a|\le K,|b|\le K} |n^{-3/2}\sum_{i,j=1}^{n^2} A_{ij2}(x,a,b)| \\
&\le \sup_{|x-y|\le K\max_{1\le j\le n}|W_j|} |n^{-1/2}\sum_{j=1}^{n}\{I(\delta_i < x) - F(x) + F(y) - I(\delta_i < y)\}|,
\end{aligned}$$

which goes to 0 by Lemma 2.4. Further, we have

$$\sup_{x,|a|\le K,|b|\le K} |n^{-3/2}\sum_{i,j=1}^{n^2} A_{ij3}(x,a,b)| = o_P(1)$$

by Lemma 2.2, and

$$\sup_{x,|a|\le K,|b|\le K} |n^{-3/2}\sum_{i,j=1}^{n^2} A_{ij4}(x,a,b)| = o_P(1)$$

by Lemma 2.1. Therefore, $\hat{p}_n(\hat{\xi}_n) - H(\hat{\xi}_n) = O_P(n^{-1/2})$. Since $H(\hat{\xi}_n) - p_n = O_P(n^{-1/2})$, the theorem is established. □

REFERENCES

(1) Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

(2) Birnbaum, Z. W. (1956). On a use of the mann-Whitney statistic. *Proceedings of the 3rd Berkeley Symposium; Mathematics, Statististics and Probability* , **1**, 13-17.

(3) Boldin, M. V. (1982). Estimation of the distribution of noise in an autoregressive scheme, *Theorey of Probability and its Applications* , **27**, 886-871.

(4) Govindarajuluu, Z. (1968). Distribution-free confidence bounds for $P(X < Y)$, *Annals of Institute of Statistical Mathematics*, **20**, 229-238.

(5) Gutmann, I., Johnson, R. A., Battachryya, G. K. and Reiser, B. (1988). Confidence limits for stress-strength models with explanatory variables, *Technometrics*, **30**, 161-168.

(6) Hannan, E. J. and Diestler (1988). *The Statistical Theory of Linear Systems*, Wiley, New York.

(7) Helperin, M. , Gilbert, P. R. and Larchin, J. M. (1987). Distribution-free confidence intervals for $Pr(X < Y)$, *Biometrics*, **43**, 71-80.

(8) Lee, S. and Wei, C. Z. (1996). On residual empirical processes of stochastic regression models with applications to time series, *Submitted for publication.*

(9) Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics,* Wiley, New York.