Equivariance, Shrinkage, and Intransitivity for Pitman Domination

by Seongmo Yoo

Pitman-Intransitive triples of estimators are constructed, based on prior work on equivariant estimators and prior work on shrinkage.

I. Introduction

According to Pitman [1], an estimator X is closer than an estimator Y to a scalar parameter θ (or, in the terminology used below, X Pitman-dominates Y) if

$$\Pr_\theta(|X-\theta|<|Y-\theta|) > 1/2 \ , \ \forall\theta.$$

This criterion is now called the Pitman Closeness Criterion (PCC). It is generalized (see Rao et al. [2] and Sen et al. [3]) with respect to the loss function $L(\cdot, \cdot)$ as follows : Let X and Y, with joint density depending on a parameter vector $\theta \in \Theta$, be estimators of θ. X is closer than Y to θ with respect to the loss function $L(\cdot, \cdot)$ (or, in the notation used below, $X \rhd Y$) if

$$\Pr_\theta(L(X,\theta) < L(Y,\theta)) > 1/2 \ , \forall\theta \in \Theta . \qquad (1)$$

This criterion is now called the Generalized Pitman Closeness Criterion(GPCC).

Pitman suggested that a median-unbiased estimator depending on a minimal sufficient statistic is well suited to PCC in the sense that it is Pitman-closest possible, and gave the "comparison theorem" for identifying classes of estimators Pitman-dominated by a minimal sufficient median-unbiased estimator. Ghosh and Sen [4], interpreting Pitman's comparison theorem in terms of Basu's theorem [5], showed that a median-unbiased estimator dominates every other estimator within the class of equivariant estimators. Nayak [6] obtained the best equivariant estimators in the sense of GPCC by using decision theoretic approaches. Kubokawa [7] showed that an estimator is median-unbiased if and only if it is the best equivariant estimator in the Pitman sense. These investigations are in a sense supportive of Pitman's idea.

Following a different line of research based on certain shrinkage constructions, David and Salem [8] considered estimating the median of asymmetric density supported by the real line. They exhibited a certain class of continuous increasing functions of such an observation, Pitman-dominating the observation itself. Robert, Hwang, and Strawderman [9] studied shrinkage constructions for Pitman-dominating an estimator of a location parameter of a density from location family. Yoo and David [10] studied shrinkage constructions for dominating an estimator of an arbitrary parameter of a not-necessarily symmetric density under PCC. Yoo and David [11] further extended these results. This direction of research is less supportive of Pitman's idea.

Pitman [1] noted that the definition of PCC is of itself intransitive: Although X Pitman-dominates Y, and Y Pitman-dominates Z, yet Z Pitman-dominates X. After that Blyth [12] found examples for discrete cases where intransitivity is manifest in the case of a single unknown parameter. David and Salem [8] constructed intransitive triples of estimators of a Laplace location parameter, each member of the triple dominating a single observation. Robert, Hwang, and Strawderman [9] constructed intransitive triples of estimators of a scale parameter of a uniform density.

The present paper is devoted to constructing a class of Pitman-intransitive triples X, $M(X)$, $T(X)$ with $X \rhd M(X)$, $T(X) \rhd X$, and $T(X) \not\rhd M(X)$ for the location parameter case and scale parameter case. The discussion is in terms of GPCC, involving loss functions of the form $L(x, \theta) = h(x-\theta)$, where $h(y) = r(y)$ on $[0, +\infty)$ and $s(-y)$ on $(-\infty, 0]$, with $r(y)$ and $s(y)$ continuous increasing on $[0, +\infty)$ and $r(0) = s(0) = 0$.

II. Preliminary Result

Let a_1 and a_2 satisfy $-\infty \le a_1 < a_2 \le +\infty$, and let I be the interval (a_1, a_2). Let X, with density $f(x;\theta)$ supported by I, have median θ. Further, for given $c \in I$, let $\lambda(\cdot)$ be any continuous increasing function (which clearly always exists, as in Yoo and David [10]) satisfying

$$\int_{\lambda(\theta)}^{c} f(x;\theta)\,dx \le 1/2, \quad a_1 < \theta \le c \qquad (2a)$$

$$\int_{c}^{\lambda(\theta)} f(x;\theta)\,dx \le 1/2, \quad c \le \theta < a_2 \qquad (2b)$$

$$\lambda(\theta) < \theta, \quad a_1 < \theta < c, \qquad (2c)$$

$$\lambda(c) = c, \qquad (2d)$$

$$\lambda(\theta) > \theta, \quad c < \theta < a_2. \qquad (2e)$$

From Yoo and David [10] we have the following lemma.

Lemma 1. Let X, median-unbiased for θ, $\theta \in I$, have density $f(x;\theta)$ with support I for all θ. Then, for $c \in (a_1, a_2)$, any estimator $T(X)$ of θ with $T(\cdot)$ continuous and

$$x < T(x) < \lambda^{-1}(x), \quad a_1 < x < c, \qquad (3a)$$

$$T(c) = c, \qquad (3b)$$

$$\lambda^{-1}(x) < T(x) < x, \quad c < x < a_2, \qquad (3c)$$

Pitman-dominates the estimator X with respect to any loss function of type L; in other words, for $\theta \in (a_1, a_2)$,

$$\Pr_\theta \left(L(T(X), \theta) < L(X, \theta) \right) > 1/2. \qquad (4)$$

III. Triples of Intransitive Estimators for Location Parameters

Lemma 1 with $c = 0$, $a_1 = -\infty$ and $a_2 = +\infty$, and the results of Ghosh and Sen [4], Nayak [6], and Kubokawa [7], lead to the following theorem, whose geometric basis is illustrated by Figure 1.

Theorem 1. *Let X, median-unbiased for a location parameter* θ , $\theta \in I = (-\infty\ , +\infty)$, *have density* $f(x\,;\theta)$ *with support I for all* θ. *Let, for any fixed* $\theta_o \in (0\ ,+\infty)$, $M(X)$ *be any estimator of* θ *of the form* $M(X) = X+b$, $b < 0$, *satisfying, for* $\theta_o \le x \le \lambda(\theta_o)$,

$$\max(x+\lambda^{-1}(\theta_o)-\theta_o\ , \lambda^{-1}(x),\ x+\theta_o-\lambda(\theta_o)) < M(x) < x\ . \qquad (5)$$

Let, for given $x_o \in (0\ , \theta_o\ )$, $T(X)$ *be any estimator of* θ *which is continuous and*

$$x < T(x) < \lambda^{-1}(x)\ ,\ x < 0 \qquad (6a)$$

$$T(0) = 0\ , \qquad (6b)$$

$$\max(\lambda^{-1}(x), M(x)) < T(x) < x\ ,\quad 0 < x < x_o \qquad (6c)$$

$$T(x_o) = M(x_o)\ , \qquad (6d)$$

$$\lambda^{-1}(x) < T(x) < M(x)\ ,\quad x_o < x < M^{-1}(\theta_o) \qquad (6e)$$

$$T(x) = M(x)\ ,\ x = M^{-1}(\theta_o) \qquad (6f)$$

$$M(x) < T(x) < x\ , M^{-1}(\theta_o) < x \le \lambda(\theta_o) \qquad (6g)$$

$$\max(\lambda^{-1}(x)\ , M(x)) < T(x) < x\ ,\quad x > \lambda(\theta_o)\ . \qquad (6h)$$

Then $X \rhd M(X)$, $T(X) \rhd X$, and $T(X) \not\rhd M(X)$; *in other words,*

$$\Pr_\theta(L(X, \theta) < L(M(X),\theta)) > 1/2\ ,\quad \forall\theta \in I\ , \qquad (7a)$$

$$\Pr_\theta(L(T(X),\theta) < L(X\ , \theta)) > 1/2\ ,\quad \forall\theta \in I\ , \qquad (7b)$$

$$\Pr_\theta(L(M(X), \theta) < L(T(X),\theta)) > 1/2\ ,\quad \theta = \theta_o\ . \qquad (7c)$$

Proof. It is easily noted that $M(X)$ is an equivariant estimator. Essentially as in Ghosh and Sen [4], Nayak [6], and Kubokawa [7], $X \rhd M(X)$. Also it is true that $T(X) \rhd X$ since $T(X)$ satisfying (6) also satisfies condition (3) in Lemma 1.

It remains to show that (7c) holds. In view of (5), it is observed that

$$\max(\lambda^{-1}(\theta_o)\ ,\ 2\theta_o - \lambda(\theta_o)) < M(\theta_o) < \theta_o \qquad (8)$$

and

$$\max(\lambda(\theta_o) + \lambda^{-1}(\theta_o) - \theta_o\ , \theta_o) < M(\lambda(\theta_o)) < \lambda(\theta_o)\ . \qquad (9)$$

(8) implies

$$\theta_o < M^{-1}(\theta_o) \qquad (10)$$

and (9) implies

$$M^{-1}(\theta_o) < \lambda(\theta_o)\ . \qquad (11)$$

Hence, in view of (10) and (11), it is true that

$$\theta_o < M^{-1}(\theta_o) < \lambda(\theta_o)\ . \qquad (12)$$

For $x_o < x < M^{-1}(\theta_o)$, it is clear that $M(x) < \theta_o$, hence, in view of (6e), we have

$$L(M(x), \theta_o) < L(T(x), \theta_o) . \qquad (13)$$

For $M^{-1}(\theta_o) < x \le \lambda(\theta_o)$, it is clear that $\theta_o <$ M(x), hence, in view of (6g), we have

$$L(M(x), \theta_o) < L(T(x), \theta_o) . \qquad (14)$$

For $x > \lambda(\theta_o)$, it is clear that $\lambda^{-1}(x) > \theta_o$ and, in view of (11), it is true that

$$M(x) > M(\lambda(\theta_o)) > M(M^{-1}(\theta_o)) = \theta_o \ ,$$

hence,

$$\min(M(x), \lambda^{-1}(x)) > \theta_o . \qquad (15)$$

Thus, in view of (6h) and (15), we have

$$L(M(x), \theta_o) < L(T(x), \theta_o) . \qquad (16)$$

Therefore, in view of (6f), (13), (14), and (16),

$$x > x_o$$

implies

$$L(M(x), \theta_o) \le L(T(x), \theta_o)$$

with equality holding only when $x = M^{-1}(\theta_o)$.

Finally, we have, for $\theta = \theta_o \in (0, +\infty)$,

$$\begin{aligned} &\Pr_\theta (L(M(X), \theta) < L(T(X), \theta)) \\ &= \Pr_\theta (X > x_o) \\ &> \Pr_\theta (X > \theta) = 1/2 . \end{aligned}$$

Remark 1. For any fixed $\theta_o \in (-\infty, 0)$, a similar argument holds by defining $M(X)$ to be any estimator of θ of the form $M(X) = X + b$, $b > 0$, satisfying, for $\lambda(\theta_o) \le x \le \theta_o$,

$$\begin{aligned} x < M(x) < \min(&x + \lambda^{-1}(\theta_o) - \theta_o , \\ &\lambda^{-1}(x),\ x + \theta_o - \lambda(\theta_o)) . \end{aligned}$$

Remark 2. In keeping with the fact that shrinkage need not be constructed with $c = 0$, the theorem also applies with arbitrary c .

Example 1. Let X be a single observation from the density

$$f(x;\theta) = 2^{-1} e^{-|x-\theta|} , \ x \in (-\infty, +\infty) ,$$

where θ is real valued unknown location parameter. It is noted that X is a median-unbiased estimator of θ. Without loss of generality, we assume $\theta > 0$. It is easy to see that $\lambda^*(\theta) = \theta - \ln(1 - e^{-\theta})$ where $\lambda^*(\theta)$ satisfies

$$\int_c^{\lambda^*(\theta)} f(x;\theta)dx = 1/2 , \ 0 < \theta < +\infty.$$

It is easy to see that $\lambda^*(x)$ is convex and has a unique minimum value 2ln2 at $x = \ln 2$. Now take $\lambda(\theta)$ satisfying

$$\lambda(\theta) = \theta + \ln(1-e^{\theta}) , \ \theta < -\ln 2$$

$$\lambda(\theta) = 2\theta , \ -\ln 2 \le \theta \le \ln 2$$

$$\lambda(\theta) = \theta - \ln(1-e^{-\theta}) , \ \theta > \ln 2 .$$

Then $\lambda(\theta)$ satisfies (2). Now take $M(X)$ and $T(X)$ satisfying (5) and (6), respectively. Then, in view of Theorem 1, $X \triangleright M(X)$, $T(X) \triangleright X$, and $T(X) \not\triangleright M(X)$.

Example 2. Let $Y_1, \cdots, Y_n$ be iid random variables having the normal density with unknown mean θ and unit variance. Let X be the sample mean(i.e., $X \equiv \overline{Y} = \sum_{i=1}^{n} Y_i / n$). It is noted that X is a median-unbiased estimator of θ. In view of Lemma 1, we can construct a function $\lambda^{-1}(x)$ satisfying (2) in the following way:

$$\int_{\lambda(\theta)}^{c} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right)dx = 1/2, \quad -\infty < \theta \le c ,$$

$$\int_{c}^{\lambda(\theta)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right)dx = 1/2, \quad c \le \theta < +\infty ,$$

$$\lambda(\theta) < \theta \ , \ -\infty < \theta < c \ ,$$
$$\lambda(c) = c \ ,$$
$$\lambda(\theta) > \theta \ , \ c < \theta < +\infty \ .$$

Now take $M(X)$, $T(X)$ satisfying (5) and (6), respectively. Then, in view of Theorem 1, $X \rhd M(X)$, $T(X) \rhd X$, and $T(X) \not\rhd M(X)$. An example of triples of intransitive estimators of θ with $c = 0$ is displayed in Figure 1 by applying Theorem 1.

IV. Triples of Intransitive Estimators for Scale Parameters

Lemma 1 with $c = 1$, $a_1 = 0$ and $a_2 = +\infty$, and the results of Ghosh and Sen [4], Nayak [6], and Kubokawa [7], lead to the following theorem, whose geometric basis is illustrated by Figure 2.

Theorem 2. *Let X, median-unbiased for a scale parameter* σ, $\sigma \in I = (0, +\infty)$, *have density* $f(x; \sigma)$ *with support I for all* σ. Let, *for any fixed* $\sigma_0 \in$

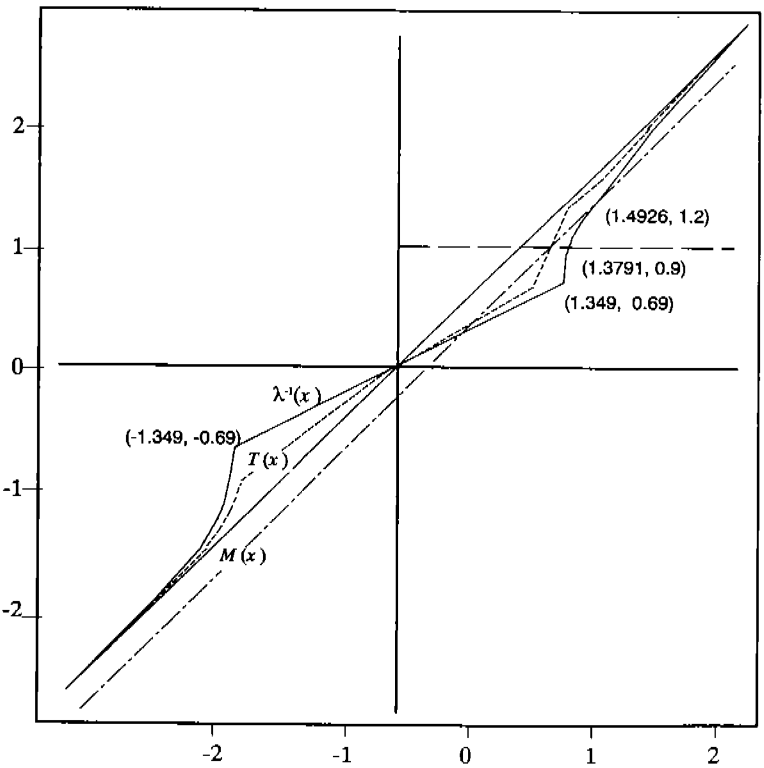

Fig. 1. Triples of intransitive estimators of θ for $N(\theta, 1)$ with $c = 0$

$(0, +\infty)$, *$M(X)$ be any estimator of σ of the form $M(X) = bX$, $0 < b < 1$, satisfying, for $\sigma_o \le x \le \lambda(\sigma_o)$,]*

$$\max(\lambda^{-1}(\sigma_o)x/\sigma_o, \lambda^{-1}(x), \sigma_o x/\lambda(\sigma_o)) < M(x) < x . \quad (17)$$

Let, for given $x_o \in (c, \sigma_o)$, $T(x)$ be any estimator of σ which is continuous and

$$x < T(x) < \lambda^{-1}(x), \quad 0 < x < c \quad (18a)$$

$$T(c) = c, \quad (18b)$$

$$\max(\lambda^{-1}(x), M(x)) < T(x) < x, \quad c < x < x_o \quad (18c)$$

$$T(x_o) = x_o, \quad (18d)$$

$$\lambda^{-1}(x) < T(x) < M(x), \quad x_o < x < M^{-1}(\sigma_o) \quad (18e)$$

$$T(x) = M(x), \quad x = M^{-1}(\sigma_o) \quad (18f)$$

$$M(x) < T(x) < x, \quad M^{-1}(\sigma_o) < x \le \lambda(\sigma_o) \quad (18g)$$

$$\max(\lambda^{-1}(x), M(x)) < T(x) < x, \quad x > \lambda(\sigma_o) . \quad (18h)$$

Then $X \rhd M(X)$, $T(X) \rhd X$, and $T(X) \not\rhd M(X)$; in other words,

$$\Pr_\sigma(L(X,\sigma) < L(M(X),\sigma)) > 1/2, \quad \forall\sigma \in I,$$
$$\Pr_\sigma(L(T(X),\sigma) < L(X,\sigma)) > 1/2, \quad \forall\sigma \in I,$$
$$\Pr_\sigma(L(M(X),\sigma) < L(T(X),\sigma)) > 1/2, \quad \sigma = \sigma_o .$$

It is also easily noted that $M(X)$ is an equivariant estimator and the proof is omitted since it essentially follows the proof of Theorem 1.

Remark 3. For any fixed $\sigma_o \in (0, c)$, a similar argument holds by defining $M(X)$ to be any estimator of σ of the form $M(X) = bX$, $b > 1$, satisfying, for $\lambda(\sigma_o) \le x \le \sigma_o$,

$$x < M(x) < \min(\lambda^{-1}(\sigma_o)x/\sigma_o, \lambda^{-1}(x), \sigma_o x/\lambda(\sigma_o)) .$$

Example 3. Let $Y_1, \cdots, Y_n$ be iid random variables having the normal density with unknown mean θ and unknown variance σ^2. Let S^2 be the mean-unbiased estimator of σ^2 (i.e., $S^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2/(n-1)$). It is observed that

$$\Pr_\sigma(0 < S^2 < \sigma^2) = \Pr_\sigma(0 < \frac{(n-1)S^2}{\sigma^2} < n-1) > 1/2 .$$

Thus, there exist k_n, $0 < k_n < 1$, such that

$$\Pr_\sigma(0 < S^2/k_n < \sigma^2) = 1/2 .$$

Therefore, $S^2/k_n \equiv X$ is a median-unbiased estimator of σ^2. Now it is clear that, for any $c > 0$, there exist a continuous increasing function $\lambda(\sigma^2)$ such that

$$\Pr_\sigma(c < X < \lambda(\sigma^2)) = 1/2, \quad \lambda(\sigma^2) > \sigma^2 > c,$$
$$\lambda(c) = c,$$
$$\Pr_\sigma(\lambda(\sigma^2) < X < c) = 1/2, \quad \lambda(\sigma^2) < \sigma^2 < c .$$

Now take $M(X)$, $T(X)$ satisfying (17) and (18), respectively. Then, in view of Theorem 2, $X \rhd M(X)$, $T(X) \rhd X$, and $T(X) \not\rhd M(X)$. An example of triples of intransitive estimators of σ^2 with $c = 1$ is displayed in Figure 2 by applying Theorem 2.

Acknowledgement

The author wishes to thank his advisor H.T. David for his helpful comments.

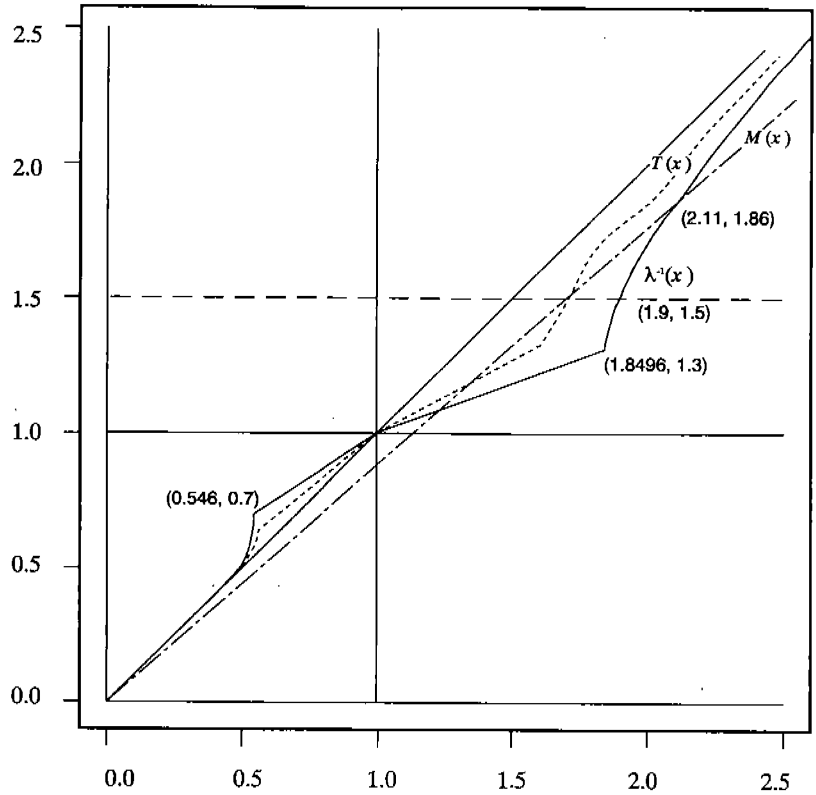

Fig. 2. Triples of intransitive estimators of σ^2 for $N(\theta, \sigma^2)$ when $n = 11$ with $c = 1$

References

[1] Pitman, E.J.G. "The "closest" estimates of statistical parameters," *Proceedings of the Cambridge Philosophical Society*, Vol. 33, pp.212-222, 1937.

[2] Rao, C.R., Keating, J.P., and Mason, R.L. "The Pitman nearness criterion and its determination," *Communications in Statistics*, Vol. A15, pp. 3173-3191, 1986.

[3] Sen, P.K., Kubokawa, T., and Saleh, A.K. "The Stein paradox in the sense of the Pitman measure of closeness," *The Annals of Statistics*, Vol. 17, pp. 1375-1386, 1989.

[4] Ghosh, M. and Sen, P.K. "Median unbiasedness and Pitman closeness," *Journal of the American Statistical Association*, Vol. 84, pp.1089-1091, 1989.

[5] Basu, D. "On statistics independent of a complete sufficient statistic," *Sankhyā*, Vol. 15, pp.377-380, 1955.

[6] Nayak, T. "Estimation of location and scale parameters using generalized Pitman nearness criterion," *Journal of Statistical Planning and Inference*, Vol. 24, pp.259-268, 1990.

[7] Kubokawa, T. "Equivariant estimation under

the Pitman closeness criterion," *Communications in Statistics*, Vol. A20, pp.3499-3523, 1991.

[8] David, H.T. and Salem, A. Shawki. "Three shrinkage constructions for Pitman-closeness in the one-dimensional location case," *Communications in Statistics*, Vol. A20, pp.3605-3627, 1991.

[9] Robert, A.P., Hwang, J.T., and Strawderman, W.E. "Is Pitman closeness a reasonable criterion?" *Journal of the American Statistical Association*, Vol. 88, pp.57-63, 1993.

[10] Yoo, S. and David, H.T. "Shrinkage constructions for Pitman domination," *Sankhyā*, Series B, *In Press*.

[11] Yoo, S. and David, H.T. "Three remarks on Pitman domination," *In Review*.

[12] Blyth, C.R. "Some probability paradoxes in choice from among random alternatives," *Journal of the American Statistical Association*, Vol. 67, pp.366-373, 1972.

Seongmo Yoo received the B.S. and M.S. degrees from Korea University in 1983 and 1985, respectively, and the Ph.D. degree from Iowa State University, Ames, in 1993, all in Statistics.

Since 1985 he has been on the research staff at Electronics and Telecommunications Research Institute, Korea, where he is currently a senior member of research staff of Radio Technology Department. His research interests are in the area of general statistical methodology, Spatial Statistics, and radiowave propagation phenomena.

Dr. Yoo is a member of American Statistical Association. He received the Gamma Sigma Delta honor for outstanding academic performance in 1992 and graduate research excellence award from Iowa State University for outstanding research performance in 1993.