Inference Properties of QIC in the Selection of Covariates for Generalized Estimating Equations

Robert N. Nyabwanga, Fredrick Onyango and Edgar O. Otumba

Department of Statistics and Actuarial Science
Maseno University, P.O. Box 333, Maseno, Kenya

This article is distributed under the Creative Commons by-nc-nd Attribution License.
Copyright © 2019 Hikari Ltd.

Abstract

The Quasi-likelihood information criterion (QIC) which results from utilizing Kullback’s I-divergence as the targeted discrepancy is widely used in the GEE framework to select the best correlation structure and the best subset of predictors. We investigated the inference properties of QIC in variable selection with focus on its consistency, sensitivity and sparsity. We established through numerical simulations that QIC had high sensitivity but low sparsity. Its type I error rate was approximately 30% which implied fairly high chances of selecting over-fit models. On the other side, it had low under-fitting probabilities. The statistical power of QIC was established to be high hence rejecting any given false null hypothesis is essentially guaranteed for sufficiently large $N$ even if the effect size is small.

Mathematics Subject Classification: 62J12, 62F07, 62F15

Keywords: Quasi-Likelihood Information Criteria, Generalized Estimating Equations, Consistency, Sparsity, Sensitivity

1 Introduction

Let $f_0(y|\theta_0)$ be the true generating model, $f(y|\theta_k)$ be the candidate model and $f(y|\hat{\theta}_k)$ be the fitted model. Further if we let $\Psi$ be the family of candidate models, then model selection seeks to search among a collection of classes $\Psi = \{\Psi(K_1)\ldots\Psi(K_L)\}$ for the fitted model $f(y|\hat{\theta}_k)$, $k \in \{1\ldots K_L\}$ which serves as the best approximation of $f_0(y|\theta_0)$. 
Akaike [2], observed that the model selected should be generalizable, a
good-fit and parsimonious. According to Burnham and Anderson [6] and Kon-
ishi and Kitagawa [7] striving for generalizability is one of the main model
selection objectives since a generalizable model will be capable of predicting
future observations with a high degree of certainty. The goodness-of-fit prin-
ciple requires that the fitted model conforms to the data used to construct
it while the principle of parsimony requires that the simplest model that ade-
quately fits the data be preferred. However, model selection should strike
a balance between goodness-of-fit and parsimony (Koniski and Kitagwa [7]).
Burnham and Anderson [6] further posit that under-fitting and over-fitting are
pertinent in determining the quality of a model. Under-fit models may lead
to biased estimates and poor predictive performance while over-fit models will
lead to results with high variability.

Fan and Li [3] observed that a good model selection criteria should be
asymptotically consistent i.e. should identify the correct model asymptoti-
cally with probability one provided the correct model is included in the set
of candidate models. Dziak [4] observes that, for consistent model selection,
two properties are required: sensitivity and sparsity. Sensitivity implies that
the model selection criteria retains all variables that should be retained with a
probability approaching one while sparsity implies that the model selection cri-
teria deletes all variables that should be deleted, with probability approaching
one.

2 Model Selection in Generalized Estimating
Equations

Let \( y_{it} (i=1, ..., n) \) be a sequence of binary responses taken on \( n \) subjects at
time points \( t=1, ..., m \) and \( X_{it} = (x_{i1}, ..., x_{im})^T \) be the \( m \times p \) matrix of covariates
for the \( i^{th} \) subject \( (i=1, ..., n) \). If \( y_i \) has any distribution from the exponential
family, then according to Liang and Zeger (1986) its probability density function,
or probability mass function, can be written as:

\[
f_Y(y; \theta, \phi) = \exp \left\{ \frac{yb - b(\theta)}{a(\phi)} + c(y, \phi) \right\}
\]

(1)

where \( \theta \) is the natural or canonical parameter of the distribution, \( \phi \) is the scale
or dispersion parameter and \( a(.) \), \( b(.) \) and \( c(.) \) are known functions. Considering
the first and second moments of \( y_{it} \) to be \( E(y_{it} = \mu_{it} \) and \( Var(y_{it}) = \sigma^2_{it} \),
then; \( \mu_{it} = b'(\theta_{it}) \) and \( \sigma^2_{it} = \sigma^2(\mu_{it}) = b''(\theta_{it})a(\phi) \). If \( Y_i \sim Binomial(m, \pi) \)
with \( m>0 \), then \( \phi=1, E(Y_i) = m\pi \) and \( Var(Y_i) = m\pi(1 - \pi) \). Further, if we
let \( y \mapsto \frac{y}{m} \) so that \( mY_i \sim Binomial(m, \pi) \), \( y = 0, \frac{1}{m}, ..., 1 \), then; \( b'(\theta) = \frac{e^\theta}{1 + e^\theta} \)
and \( b''(\theta)a(\phi) = \frac{e^\theta}{m(1 + e^\theta)^2} = \frac{\pi(1 - \pi)}{m} = \frac{\mu(1 - \mu)}{m} \).
Definition 2.1 Suppose that we have independent observations $Y_i (i=1,\ldots,n)$ from $n$ subjects and for each subject $i$, $m$ observations are made such that $Y_{it}$ denote the $t^{th}$ response $(t=1,2,\ldots,m)$ and $X_{it}=\{X_{it1},X_{it2},\ldots,X_{itm}\}'$ denote a $p \times 1$ vector of covariates associated with $Y_{it}$ where $(\cdot)'$ denotes transpose. Let $Y_i = [y_{i1},\ldots,y_{im}]'$ denote the response vector for the $i^{th}$ subject and $X_i = [X_{i1},\ldots,X_{im}]'$ be the $m \times p$ corresponding covariates matrix. Further, if $E(Y_{it}) = \mu_{it}$ such that $g(\mu_{it}) = \eta_{it} = X_{it}'\beta$, where $\beta = [\beta_1,\ldots,\beta_p]'$ is a $p \times 1$ vector of regression parameters and $X_{it}$ is the $i^{th}$ row of $X_i$; $\text{var}(Y_{it}/X_{it}) = \phi v(\mu_{it})$, where $v(.)$ is a known variance function of $\mu_{it}$ and $\phi$ is a scale parameter which may need to be estimated by $\hat{\phi} = \frac{1}{N-p} \sum_{i=1}^{n} \sum_{t=1}^{m} e_{it}^2$ where $N = \sum_{i=1}^{n} m$ and $p$ is covariates dimensionality and if an $m \times m$ working correlation matrix $R(\alpha)$ is assumed for each $Y_{it}$ and is assumed to be a fully specified $h \times 1$ vector of unknown parameters, $\alpha = [\alpha_1,\ldots,\alpha_h]'$ such that the corresponding working covariance matrix for $Y_{it}$ is given as $V_i = \phi A_i^{1/2} R_i(\alpha) A_i^{1/2}$ where $A_i$ is a $m \times m$ diagonal matrix with $g(\mu_{it})$ as the $t^{th}$ diagonal element and $R_i(\alpha)$ is an $m \times m$ working correlation matrix that depends on the correlation parameter $\alpha$, then, according to Liang and Zeger [5] the method of generalized estimating equations (GEE) is appropriate for modeling $Y_i$. The quasi-likelihood GEE parameter estimates of $\beta$ could be obtained by solving the following system utilizing iteratively re-weighted least squares method:

$$U(\hat{\beta},R_i,\varphi_i) = \sum_{i=1}^{n} D_i'V_i^{-1}(y_i - \mu_i) = 0 \quad (2)$$

$D_i = \frac{d\mu_i}{d\beta}$ which is the first derivative of the response mean with respect to the regression parameters. $\varphi_i \equiv (Y_i,X_i), i=1,2,\ldots,n$ indicates the data at hand. Since the GEE depend on both $\beta$ and correlation parameters $\alpha$ and have no closed-form solution, iterative two-stage estimation procedure of $\beta$ and the nuisance parameters $(\alpha$ and $\phi)$ is required. $(y_i - \mu_i)$ is a residual vector which measures deviations of observed responses of the $i^{th}$ subject from its mean.

Solving (2) yields the quasi-likelihood-based estimator of $\hat{\beta}$. Under mild regularity conditions $\sqrt{n}(\hat{\beta}_G - \beta) \to N(0,V_{LZ})$ i.e. $\hat{\beta}$ is $\sqrt{n}$-consistent for $\beta : \hat{\beta} \to \beta$ as $n \to \infty$. $V_{LZ}$ is a covariance matrix based on the sandwich estimator given by:

$$V_{LZ} = B^{-1} \hat{M}_{LZ} B^{-1} \quad (3)$$

Where $B = \frac{1}{n} \sum_{i=1}^{n} D_i'V_i^{-1}D_i$ and $\hat{M}_{LZ} = \frac{1}{n} \sum_{i=1}^{n} D_i'V_i^{-1}\text{Cov}(Y_i)V_i^{-1}D_i$. If $(V_i)$ is correctly specified, $V_{LZ}$ reduces to $\sum_{i=1}^{n} D_i'V_i^{-1}D_i$ which is referred to as model-based variance estimator [1].
2.1 Quasi-Likelihood Information Criteria (QIC) and Variable Selection in GEE

QIC was proposed by Pan [8] as a modification of AIC [2] for use in the GEE framework in the selection of the covariates for the mean structure. QIC was based on quasi-likelihood function under independent correlation structure and is Mathematically defined as

$$QIC^R = -2Q(\hat{\beta}(R); I, \varphi) + 2tr(\hat{\Omega}_I \hat{V}_r)$$  \hspace{1cm} (4)

Where $\hat{\Omega}_I$ is the model-based variance estimator under the independence working correlation structure given as:

$$\hat{\Omega}_I = E_0 \{ - \frac{d^2Q(\beta; I, \varphi)}{d\beta d\beta^T} \} \bigg|_{\beta = \beta_0} = \sum_{i=1}^{n} D_i^T V_i^{-1} D_i$$  \hspace{1cm} (5)

Where $D_i = \frac{d\mu_i}{d\beta^T}$ and

$$\hat{V}_r = \hat{\Omega}_I C \hat{\Omega}_I$$  \hspace{1cm} (6)

Where $C = \sum_{i=1}^{n} D_i^T V_i^{-1} (Y_i - \mu_i)(Y_i - \mu_i)^T V_i^{-1} D_i$ is the robust or sandwich variance estimator under the working correlation structure R and $tr(\hat{\Omega}_I \hat{V}_r)$ is the trace of the product of the two matrices i.e the sum of the diagonal elements of the matrix and is considered a measure of total variability or spread.

2.2 Inference Properties of QIC in the selection of Covariates for GEEs

To formally state some of these inference properties investigated, let $Q_n(p)$ be the quasi-likelihood of a model with p parameters based on a sample size n and $Q_n(p_0)$ be the quasi-likelihood of the model with $p_0$ correct parameters. If $p > p_0$, the model with p parameters is nested in the model with $p_0$ parameters so that $Q_n(p_0)$ is obtained by setting $p - p_0$ parameters in the larger model to constants which can be assumed to be zero without loss of generality. Models in which $p < p_0$ are mis-specified and the models with with $p \geq p_0$ are correctly specified or over-specified.

Based on QIC, the general form of the model selection criteria can be expressed in the form:

$$QIC_n(p) = -2Q_n(p) + \frac{\varphi}{n} trace(\Theta)$$  \hspace{1cm} (7)

where $\frac{\varphi}{n} = 2$ and $\Theta = \hat{\Omega}_I \hat{V}_r$. Using the general form, the model is selected that corresponds to

$$\hat{p} = \arg \min_{p \leq m} QIC_n(p)$$  \hspace{1cm} (8)
Inference properties of QIC in the selection of covariates for GEE

If \( p < p_0 \), then the model with \( p \) parameters is mis-specified so that

\[
\text{Plim}_{n \to \infty} \ln(Q_n(p)) < \text{Plim}_{n \to \infty} \ln(Q_n(p_0))
\]  \( (9) \)

Hence from (7), (9) and \( \lim_{n \to \infty} \frac{\varphi}{n} = 0 \) it follows that;

\[
\lim_{n \to \infty} \Pr[QIC_n(p_0) \geq QIC_n(p)] = \lim_{n \to \infty} \Pr[\hat{p} < p_0] \leq \sum_{\hat{p} < p_0} \lim_{n \to \infty} \Pr[QIC_n(p_0) \geq QIC_n(p)] = 0 \]  \( (10) \)

For \( p > p_0 \), it follows from the likelihood ratio test that;

\[
2(\ln(Q_n(p)) - \ln(Q_n(p_0))) \overset{d}{\to} \chi^2_{p-p_0}
\]  \( (12) \)

In the QIC case,

\[
Q_n(p_0) - Q_n(p) = 2(\ln(Q_n(p)) - \ln(Q_n(p_0)) - (p - p_0)\frac{\varphi}{n}\text{trace}(\Theta) \overset{d}{\to} X_{p-p_0} - (p - p_0)\frac{\varphi}{n}\text{trace}(\Theta)
\]  \( (13) \)

hence;

\[
\lim_{n \to \infty} \Pr[QIC_n(p_0) > QIC_n(p)] = \Pr[X_{p-p_0} > (p - p_0)\frac{\varphi}{n}\text{trace}(\Theta) > 0]
\]  \( (14) \)

Therefore it follows that:

i if \( \lim_{n \to \infty} \Pr(\hat{p} \geq p_0) \to 1 \) and \( \lim_{n \to \infty} \Pr(\hat{p} > p_0) > 0 \), then QIC will be inferred to over-fit in which case it will include all the \( p_0 \) parameters plus some spurious ones i.e. \( p_0 \subset \hat{p} \). \( \Pr(\hat{p} \geq p_0) \) is the type 1 error rate of the model selection criteria.

ii if \( \lim_{n \to \infty} \Pr(\hat{p} < p_0) = 1 \) but \( \lim_{n \to \infty} \Pr(\hat{p} < p_0) > 0 \), then QIC will be inferred to under-fit in which case it will exclude some important variables from the selected model i.e.\( p_0 \subsetneq \hat{p} \). \( \Pr(\hat{p} < p_0) \) is the type II error rate (\( \beta \)). The values of \( (1-\beta) \) represents the power of the test.

iii if \( \lim_{n \to \infty} \Pr(\hat{p} = p_0) = 1 \) but \( \lim_{n \to \infty} (\hat{p} > p_0) = 0 \) and \( \lim_{n \to \infty} (\hat{p} < p_0) = 0 \), then QIC will be inferred to strongly consistent in selecting the true model i.e. it will almost surely select the true model.
iv if \( \lim_{n \to \infty} Pr(\hat{p} = p_0) \to 1 \) but \( \lim_{n \to \infty}(\hat{p} > p_0) \to 0 \) and \( \lim_{n \to \infty}(\hat{p} < p_0) \to 0 \), then QIC will be inferred to be weakly consistent in selecting the true model i.e. it will select the true model with probabilities converging to one.

### 2.3 Simulation Design

In the numerical simulation we determine the consistency, over-fitting, under-fitting, sensitivity and sparsity of QIC in selecting the true GEE model. We considered a model with four covariates \( x_1, x_2, x_3 \) and \( x_4 \). The binary response \( y_{it} \) has the conditional expectation \( \mu_{it} \):

\[
\mu_{it} = E\{y_{it}|X_{1,it}, X_{2,it}, X_{3,it}, X_{4,it}\} 
\]

(15)

\( \mu_{it} \) can be connected with the covariates through:

\[
\text{logit}(\mu_{it}) = \beta_0 + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \beta_3 X_{3,it} + \beta_4 X_{4,it} 
\]

(16)

where \( i \in \{1...n\} \) and \( t \in \{1...m\} \). \( \{\beta_0, \beta_1, \beta_2\} = \{0.25, -0.25, -0.25\} \) and \( \beta_0 = 0 \) if \( p \neq 1, 2 \). This implies that the model with \( X_{1,it} \) and \( X_{2,it} \) is the true model. \( x_{1,it} \sim N(0,1) \); \( x_{2,it} \sim \text{Bernoulli}(0.5) \) and \( \{x_3, x_4\} \sim \text{Uniform}[0,1] \).

The true correlation structure \( R_0 \) was assumed to be AR-1 \( \alpha \in (0, 0.5) \). The simulation studies were based on \( 2^k \) factorial design. The narrow model included the intercept term and \( x_1 \) and the final model is selected from the narrow model combined with the other \( 2^3 \) candidate models. The Simulation results were based on 1,000 replications and the performance of QIC was based on how many times it chose the true data generating model, type I(\( \alpha \)) and type II(\( \beta \)) error rates and the power of the test \( (1 - \beta) \).

### 3 Numerical Results On the Selection of the True Model

The results in table 1 show that QIC’s selection rates of the true model increase as the sample size increases. Likewise, increasing the number of measurements per subject and the level of within-subject correlation results to increased proportion of selection of the true model. The type I error rates which are the over-fitting probabilities of QIC are shown in Table 2 indicate that for both \( \alpha = 0.2 \) and \( \alpha = 0.5 \), \( m = 3, 6 \) and 9 and for \( n \) in the range, \( 20 \leq n \leq 200 \), the range of type I error rate is \( \approx 30\% \). However, it is noticeable that the type I error rate is higher for smaller sample sizes and seems to diminish as the the sample size increases but seems not to approach zero. The Type I error rate of at least 30% indicates fairly high chances of QIC selecting over-fit models by wrongly including covariates whose coefficients are zero.
The high type I error rates implies reduced risk of type II error hence increased statistical power of QIC i.e. increased ability to make the correct inclusion of the important variables in the selected model. These are illustrated in table 3 and indicate that the type II error rates are about 30% for small samples, but quickly diminish to zero as the sample size increases. This implies that for small samples QIC has some chances of under-fitting at the rate of about 30%. For \( n \geq 50 \), QIC has little or no chances of under-fitting. The power test results show that the power of the test increases with \( n \), so that rejecting any given false null hypothesis is essentially guaranteed for sufficiently large \( n \) even if the effect size is small. This makes QIC good in predictive modeling.

4 Theoretical Results

Proposition 4.1 Let \( M_c = \{m_1, m_2, ..., m_p\} \) be the set of all \( p \) candidate models. We can partition \( M_c \) into two sets: \( M_+ \) set of over-specified models i.e. candidate models that include the true model, i.e. \( M_+ = \{m \in M_c/m_\ast \subset m\} \) and \( M_- = M_c \setminus (M_+) \), the set of under-specified models. If we let \( \overline{m} \) denote the model selected by QIC then;

i. \( \Pr(\overline{m} = M_+) > 0 \). i.e. the probability of QIC selecting over-fit models is greater than zero.

ii. \( \Pr(\overline{m} = M_+) - \Pr(\overline{m} = M_-) > 0 \) i.e. Over-fitting probability is greater than the under-fitting probability.

iii. \( \Pr(\overline{m} = M_-) \to 0 \) as \( n \to \infty \).

iv. \( \Pr(\overline{m} = m_\ast) \to 1 \) as \( n \to \infty \). This implies that, with probability approaching one, the QIC procedure selects the correct data generating model. However, this consistency is dependent on \( \Pr(\overline{m} = M_+) \) i.e. \( \Pr\{\Pr(\overline{m} = M_+) \to 0 \mid \Pr(\overline{m} = M_+) \to 0 \} \) as \( n \to \infty \).

Proposition 4.2 If we partition \( \beta_\ast \) the true value of \( \beta \) into truly non-zero(NZ) and truly zero (Z) coefficients as follows: \( \text{NZ} = \{j : \beta_j \neq 0\} \) and \( Z = \{j : \beta_j = 0\} \). Further, if we let \( \beta_{NZ} \) denote the vector of non-zero coefficients and \( \beta_Z \) denote the vector of zero coefficients. Then;

i. \( \Pr \{\exists j \in NZ : \hat{\beta}_j = 0\} = o(1) \) i.e. there exists \( \hat{\beta}_{NZ} \), a sequence of solutions of GEE model such that non-zero coefficients are included in the model selected by QIC with probability approaching one(sensitivity).

ii. \( \Pr(\beta_Z = \beta_{*Z}) \to 1 \) as \( n \to \infty \) i.e. QIC is not sparse.
5 Conclusion

We established that QIC had high sensitivity i.e. it included all the important variables in the final model selected but had low sparsity i.e. did not delete all the non-important variables with probability one. Hence, the study established that QIC had some realistic chances of selecting over-fit models. This is also attributed to the high type I error rate of about 30%. Likewise, QIC had a high statistical power which resulted from the low type II error rate hence rejecting any given false null hypothesis is essentially guaranteed for sufficiently large samples even if the effect size is small making QIC good in predictive modeling.

References


Received: April 15, 2019; Published: June 10, 2019
Table 1: Proportion of True Model Selection.

<table>
<thead>
<tr>
<th>m</th>
<th>20</th>
<th>30</th>
<th>50</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>m=3</td>
<td>30.8</td>
<td>42.3</td>
<td>55.2</td>
<td>66.3</td>
<td>66.5</td>
</tr>
<tr>
<td>α = 0.2</td>
<td>m=6</td>
<td>37.6</td>
<td>51.7</td>
<td>62.8</td>
<td>64.2</td>
</tr>
<tr>
<td></td>
<td>m=9</td>
<td>41.2</td>
<td>53.4</td>
<td>64.7</td>
<td>68.0</td>
</tr>
<tr>
<td>m=3</td>
<td>36.0</td>
<td>46.3</td>
<td>55.5</td>
<td>66.3</td>
<td>70.5</td>
</tr>
<tr>
<td>α = 0.5</td>
<td>m=6</td>
<td>41.7</td>
<td>53.5</td>
<td>58.3</td>
<td>68.5</td>
</tr>
<tr>
<td></td>
<td>m=9</td>
<td>46.2</td>
<td>57.3</td>
<td>64.0</td>
<td>70.3</td>
</tr>
</tbody>
</table>

Table 2: Model selection summary by QIC. Type I Error Rate.

<table>
<thead>
<tr>
<th>m</th>
<th>20</th>
<th>30</th>
<th>50</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>m=3</td>
<td>α = 0.2</td>
<td>0.345</td>
<td>0.325</td>
<td>0.336</td>
<td>0.328</td>
</tr>
<tr>
<td>m=6</td>
<td>0.428</td>
<td>0.362</td>
<td>0.346</td>
<td>0.358</td>
<td>0.320</td>
</tr>
<tr>
<td>m=9</td>
<td>0.465</td>
<td>0.418</td>
<td>0.344</td>
<td>0.320</td>
<td>0.323</td>
</tr>
<tr>
<td>m=3</td>
<td>α = 0.5</td>
<td>0.465</td>
<td>0.418</td>
<td>0.344</td>
<td>0.320</td>
</tr>
<tr>
<td>m=6</td>
<td>0.369</td>
<td>0.343</td>
<td>0.377</td>
<td>0.330</td>
<td>0.295</td>
</tr>
<tr>
<td>m=9</td>
<td>0.486</td>
<td>0.414</td>
<td>0.358</td>
<td>0.297</td>
<td>0.278</td>
</tr>
</tbody>
</table>

Table 3: Model selection summary by QIC. Type II Error and Statistical Power.

<table>
<thead>
<tr>
<th>m</th>
<th>20</th>
<th>30</th>
<th>50</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>m=3</td>
<td>β</td>
<td>0.34</td>
<td>0.25</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>1-β</td>
<td>0.66</td>
<td>0.75</td>
<td>0.89</td>
<td>0.99</td>
</tr>
<tr>
<td>α = 0.2</td>
<td>m=6</td>
<td>β</td>
<td>0.20</td>
<td>0.12</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>1-β</td>
<td>0.80</td>
<td>0.88</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td>m=9</td>
<td>β</td>
<td>0.12</td>
<td>0.05</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1-β</td>
<td>0.88</td>
<td>0.95</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>m=3</td>
<td>β</td>
<td>0.27</td>
<td>0.19</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>1-β</td>
<td>0.73</td>
<td>0.81</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>α = 0.5</td>
<td>m=6</td>
<td>β</td>
<td>0.13</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>1-β</td>
<td>0.87</td>
<td>0.96</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>m=9</td>
<td>β</td>
<td>0.05</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1-β</td>
<td>0.95</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>