10.2478/v10048-009-0003-9

Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model with Normal Errors

Martina Chvosteková, Viktor Witkovský

Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 84104 Bratislava, Slovakia
E-mail: chvosta@gmail.com, witkovsky@savba.sk

In this paper we present an exact likelihood ratio test (LRT) for testing the simple null hypothesis on all parameters of the linear regression model with normally distributed errors. In particular, we consider the simultaneous test for the regression parameters, β, and the error standard deviation, σ. The critical values of the LRT are presented for small sample sizes and a small number of explanatory variables for usual significance levels, $\alpha = 0.1$, 0.05, and 0.01. The test is directly applicable for construction of the $(1-\alpha)$-confidence region for the parameters (β, σ) and the simultaneous tolerance intervals for future observations in linear regression models. For comparison, the suggested method for construction of the tolerance factors of the symmetric $(1-\gamma)$-content simultaneous $(1-\alpha)$-tolerance intervals is illustrated by a simple numerical example.

Keywords: linear regression model; exact likelihood ratio test; simultaneous tolerance intervals; confidence-set approach.

1. Introduction

In the paper we present an exact likelihood ratio test for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$ on the parameters β and σ of the linear regression model $Y = X\beta + \sigma Z$ with normally distributed errors, $Z \sim N(0, I_n)$.

Although the derivation of the exact distribution of the likelihood ratio test (LRT) statistic under the null hypothesis H_0 is straightforward, it seems that the result is not available in the standard (statistical) literature on linear regression models nor in the literature on their applications in measurement science and metrology.

The test is directly applicable for construction of the $(1-\alpha)$-confidence region for the parameters (β, σ) and for construction of the simultaneous tolerance intervals for future observations in linear regression models. The simultaneous tolerance intervals are important for many measurement procedures. The most common application for simultaneous tolerance intervals is the multiple-use calibration problem; see e.g. [7], [6], [1]. The tolerance intervals have been recognized and considered in various settings by many authors, see e.g. [9], [8], [10], [3], [4], [5], [6], and [2]. These simultaneous tolerance intervals are constructed such that with given confidence level $(1-\alpha)$ at least a specified proportion $(1-\gamma)$ of the population is contained in the tolerance interval for all possible values of the predictor variates. All known simultaneous tolerance intervals in regression are conservative in that the actual confidence level exceeds the nominal level $(1-\alpha)$.

Here we present the method for computing the critical values of the LRT as well as the tables of the critical values for selected small sample sizes in the range $n = k+1, \ldots, 100$ with k explanatory variables, $k = 1, \ldots, 10$, and usual significance levels $\alpha = 0.1, 0.05$, and 0.01.

The suggested method for construction of the simultaneous tolerance intervals and/or the tolerance factors of the symmetric $(1-\gamma)$-content simultaneous $(1-\alpha)$-tolerance intervals is illustrated by a simple numerical example.

2. Likelihood ratio test of the hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$

Consider the linear regression model $Y = X\beta + \sigma Z$ with normally distributed errors, where Y represents the n-dimensional random vector of response variables, X is the $n \times k$ matrix of non-stochastic explanatory variables (for simplicity, here we assume that X is a full-rank matrix), β is a k-dimensional vector of regression parameters, Z is an n-dimensional vector of standard normal errors, i.e. $Z \sim N(0, I_n)$, and σ is the error standard deviation, $\sigma > 0$.

Here we consider the LRT for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$. Based on the above assumptions the log-likelihood function, denoted as $\ell(\beta, \sigma \,|\, Y = y, X)$, is given by

$$\begin{aligned} \ell(\beta, \sigma \,|\, y, X) = & -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) \\ & - \frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta). \end{aligned} \quad (1)$$

The (-2)-multiple of the logarithm of the LRT statistic, say $\lambda(y \,|\, X)$ for observed value y of Y, and given X, for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ is given by

$$\begin{aligned} \lambda(y \,|\, X) &= -2\left(\sup_{(\beta,\sigma)\in H_0} \ell(\beta, \sigma \,|\, y, X) - \sup_{(\beta,\sigma)} \ell(\beta, \sigma \,|\, y, X)\right) \\ &= -2\left(\ell(\beta_0, \sigma_0 \,|\, y, X) - \ell(\hat{\beta}_{ML}, \hat{\sigma}_{ML} \,|\, y, X)\right) \\ &= \frac{1}{\sigma_0^2}(y - X\beta_0)'(y - X\beta_0) - n\log\left(\frac{\hat{\sigma}^2_{ML}}{\sigma_0^2}\right) - n, \end{aligned} \quad (2)$$

where $\hat{\beta}_{ML} = \hat{\beta} = (X'X)^{-1}X'y$ is the standard least squares estimate (LSE) of β (which is also the MLE of β) and $\hat{\sigma}_{ML}$ is the maximum likelihood estimate (MLE) of the standard deviation σ, i.e. $\hat{\sigma}_{ML} = \sqrt{(y - X\hat{\beta})'(y - X\hat{\beta})/n}$. Under given model assumptions, and under the null hypothesis H_0, it is straightforward to derive the distribution of the test statistic $\lambda(Y\,|\,X)$:

$$\begin{aligned}\lambda(Y\,|\,X) &\sim \frac{1}{\sigma_0^2}(Y - X\beta_0)'(Y - X\beta_0)\\ &\quad - n\log\left(\frac{(Y - X\beta_0)'M_X(Y - X\beta_0)}{n\sigma_0^2}\right) - n\\ &\sim Z'Z - n\log(Z'M_XZ) + n(\log(n) - 1)\\ &\sim Z'(P_X + M_X)Z - n\log(Z'M_XZ) + n(\log(n) - 1)\\ &\sim Q_k + Q_{n-k} - n\log(Q_{n-k}) + n(\log(n) - 1),\end{aligned} \tag{3}$$

where $P_X = X(X'X)^{-1}X'$, $M_X = I_n - P_X$, $Z \sim N(0, I_n)$, $Q_k \sim \chi^2_k$ and $Q_{n-k} \sim \chi^2_{n-k}$ are two independent random variables with chi-square distributions, with k and $n-k$ degrees of freedom, respectively. Note that the distribution of $\lambda(Y\,|\,X)$ does not depend on the particular form of the regression design matrix X but only on the number of observations n and on $k = rank(X)$ – the rank of the matrix X, i.e. $\lambda(Y\,|\,X) \sim \lambda(Y\,|\,n, k)$.

This LRT rejects the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ for large values of the observed test statistic $\lambda(y\,|\,X)$, i.e. for the given significance level $\alpha \in (0, 1)$ the test rejects the null hypothesis if

$$\lambda(y\,|\,X) > \lambda_{1-\alpha}, \tag{4}$$

where $\lambda_{1-\alpha}$ is the $(1-\alpha)$-quantile of the distribution of the random variable $\lambda(Y\,|\,n, k)$, given in (3). The quantiles $\lambda_{1-\alpha}$ could be evaluated numerically, by inverting the cumulative distribution function, of the random variable $\lambda(Y\,|\,n, k)$, denoted by $\mathcal{F}_{LR}(\cdot)$:

$$\begin{aligned}\mathcal{F}_{LR}(x) &= \Pr(\lambda(Y\,|\,n, k) \le x)\\ &= \Pr(Q_k \le x - Q_{n-k} + n\log(Q_{n-k}) - n(\log(n) - 1))\\ &= \int_0^\infty \mathcal{F}_{\chi^2_k}(x - q_{n-k} + n\log(q_{n-k}) - n(\log(n) - 1))\\ &\qquad \times f_{\chi^2_{n-k}}(q_{n-k})\,\mathrm{d}q_{n-k},\end{aligned} \tag{5}$$

where $\mathcal{F}_{\chi^2_k}(\cdot)$ denotes the cumulative distribution function of the chi-square distribution with k degrees of freedom, and $f_{\chi^2_{n-k}}(\cdot)$ denotes the probability density function of the chi-square distribution with $n-k$ degrees of freedom.

Notice that since the family of normal distributions meets the regularity conditions, from the standard asymptotic result about the distribution of the LRT $\lambda(Y\,|\,X)$ we get that $\lambda_{1-\alpha} \to \chi^2_{k+1,1-\alpha}$ as $n \to \infty$, where $\chi^2_{k+1,1-\alpha}$ denotes the $(1-\alpha)$-quantile of chi-square distribution with $k+1$ degrees of freedom.

The critical values of the LRT are presented in the enclosed tables for different number of explanatory variables, $k = 1, \ldots, 10$, selected small sample sizes, $n = k+1 : (1) : 40$, $n = 45 : (5) : 100$ and ∞, and for the usual significance levels $\alpha = 0.1$ (Table 2), $\alpha = 0.05$ (Table 3), and $\alpha = 0.01$ (Table 4).

Alternatively, the test of the hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ could be based on the test statistic $F^\star$ defined as $F^\star = \lambda(Y\,|\,X)/(kS^2/\sigma_0^2)$, where $S^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})/(n-k)$ and $\hat{\beta} = (X'X)^{-1}X'Y$. Hence, we get

$$\begin{aligned}F^\star &= \frac{1}{k}\frac{(\hat{\beta} - \beta_0)'X'X(\hat{\beta} - \beta_0)}{S^2} + \frac{n-k}{k}\\ &\quad - \frac{n}{k}\frac{\log\left((n-k)S^2/n\sigma_0^2\right) + 1}{S^2/\sigma_0^2},\end{aligned} \tag{6}$$

Note that the leading term in $F^\star$ is the standard F-statistic for testing the hypothesis on regression parameters $H_0 : \beta = \beta_0$ against the alternative $H_1 : \beta \ne \beta_0$. Under null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ we get directly

$$F^\star \sim \frac{Q_k/k}{Q_{n-k}/n-k} + \frac{n-k}{k} - \frac{n}{k}\frac{\log(Q_{n-k}/n) + 1}{Q_{n-k}/n-k}. \tag{7}$$

Then, the test rejects the null hypothesis if

$$F^\star_{obs} > F^\star_{1-\alpha}, \tag{8}$$

where $F^\star_{obs}$ denotes the observed value of the statistic $F^\star$ and $F^\star_{1-\alpha}$ is the $(1-\alpha)$-quantile of the distribution of the random variable $F^\star$. Similarly as before, the quantiles $F^\star_{1-\alpha}$ could be evaluated numerically, by inverting the cumulative distribution function of the random variable $F^\star$, denoted by $\mathcal{F}_{F^\star}(x)$:

$$\begin{aligned}\mathcal{F}_{F^\star}(x) &= \Pr(F^\star \le x)\\ &= \Pr\left(Q_k \le \frac{xkQ_{n-k}}{n-k} - Q_{n-k} + n\left(\log\left(\frac{Q_{n-k}}{n}\right) + 1\right)\right)\\ &= \int_0^\infty \mathcal{F}_{\chi^2_k}\left(\frac{xkq_{n-k}}{n-k} - q_{n-k} + n\left(\log\left(\frac{q_{n-k}}{n}\right) + 1\right)\right)\\ &\qquad \times f_{\chi^2_{n-k}}(q_{n-k})\,\mathrm{d}q_{n-k}.\end{aligned} \tag{9}$$

Numerical Algorithm

The Matlab algorithm for computing the quantiles $\lambda_{1-\alpha}$ and $F^\star_{1-\alpha}$ is available upon request from the authors.

For those interested in the idea of the algorithm, we present here a simplified working version of the Matlab code used for calculation of the quantiles of the distribution of the random variable $\lambda(Y\,|\,n, k)$, which is based on the following functions available in Matlab:

`FZERO` Single-variable nonlinear zero finding. `X = FZERO(FUN,X0)` tries to find a zero of the function `FUN` near `X0`, if `X0` is a scalar.

`QUADGK` Numerically evaluate integral by adaptive Gauss-Kronrod quadrature. `Q = QUADGK(FUN,A,B)` attempts to approximate the integral of scalar-valued function `FUN` from `A` to `B` using high order global adaptive quadrature and default error tolerances.

`CHI2CDF` Chi-square cumulative distribution function. `P = CHI2CDF(X,V)` returns the chi-square cumulative distribution function with `V` degrees of freedom at the values in `X`, (Statistics Toolbox).

`CHI2PDF` Chi-square probability density function (pdf). `Y = CHI2PDF(X,V)` returns the chi-square pdf with `V` degrees of freedom at the values in `X`, (Statistics Toolbox).

Then, for example, the 0.95-quantile $\lambda_{1-0.05}$ of the random variable $\lambda(Y \mid n, k)$, based on the linear regression model with $n = 15$ observations and $k = 2$ regressors, could be calculated by calling the user defined function `LRinv`:

```
>> quantile = LRinv(2,15,0.95)
```

which gives as a result the value

```
>> quantile = 8.6813
```

The function `LRinv` could be defined as follows (note that it is calling other auxiliary functions, given bellow):

```
function quantile = LRinv(k,n,prob)
x0 = k + 1;
quantile = fzero(@(x)LRcdf(x,n,k)-prob,x0);
%
function cdf = LRcdf(x,n,k)
cdf = quadgk(@(q)LRfun(q,n,k,x),0,Inf);
%
function fun = LRfun(q,n,k,x)
fun = chi2cdf(x - n * (log(n)-1) - q ...
    + n * log(q), k) .* chi2pdf(q, n-k);
```

3. Simultaneous tolerance intervals

The suggested exact tests, (4) and (8), for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$, could be directly used to construct the exact simultaneous confidence region for all parameters of the linear regression model. In particular, based on LRT (4), the exact $(1-\alpha)$-confidence region for the parameters β and σ is given as

$$C^{LR}_{1-\alpha}(Y \mid X) = \{(\beta, \sigma) : \lambda(Y \mid X) \leq \lambda_{1-\alpha}\}. \tag{10}$$

Moreover, the confidence regions based on the exact tests could be directly used for constructing the $(1-\gamma)$-content simultaneous $(1-\alpha)$-tolerance intervals in linear regression model with normal errors. These intervals are constructed such that, with confidence $1-\alpha$, we can claim that at least a specified proportion, $1-\gamma$, of the population is contained in the tolerance interval, for all possible values of the predictor variates, see e.g. [1] and [2].

Several authors considered the confidence-set approach for constructing the simultaneous tolerance intervals in linear regression. In [10], Wilson used an ellipsoidal confidence set for the regression coefficients, β, and standard deviation of the residuals, σ, which imposes an unnecessary lower bound on σ, as noted by Limam and Thomas in [5]. They considered two alternative confidence sets. The first is a modification of Wilson's confidence set, which removes the lower bound imposed on σ, and the second is based on a product set formed from the ellipsoidal confidence set for β and a one-sided confidence interval for σ.

They found the tolerance intervals based on the product confidence set to be efficient and easy to compute compared with those constructed from the ellipsoidal and the modified ellipsoidal confidence sets.

Here we suggest a new procedure for simultaneous $(1-\alpha)$-tolerance intervals covering at least the $(1-\gamma)$-content about the true mean $x'\beta$, for any vector $x = (x_1, \ldots, x_k)'$ of explanatory variables, based on the LRT (4) and the confidence set (10), defined by

$$\begin{aligned}\mathcal{T}^{1-\gamma}_{1-\alpha}(x \mid Y, X) = \Big[&\inf_{(\beta,\sigma)\in C_{1-\alpha}(Y \mid X)} \left\{x'\beta + u_{\gamma_1}\sigma\right\};\\ &\sup_{(\beta,\sigma)\in C_{1-\alpha}(Y \mid X)} \left\{x'\beta + u_{1-\gamma_2}\sigma\right\}\Big],\end{aligned} \tag{11}$$

where u_{γ_1} and $u_{1-\gamma_2}$ are the pre-specified quantiles of the standard normal distribution, and such that $\gamma = \gamma_1 + \gamma_2$, with $\gamma \in (0, 1)$.

As a special case we get the symmetric tolerance intervals about the fitted regression function, with $\gamma_1 = \gamma_2 = \gamma/2$. Notice, that immediately we get the following probability statement

$$\begin{aligned}\Pr\Big(\Pr\Big(x'\beta + \sigma Z \in \mathcal{T}^{1-\gamma}_{1-\alpha}(x \mid Y, X)\Big) \geq 1-\gamma,&\\ \text{for all } x \text{ and } Z \sim N(0,1), Z \perp Y\Big) \geq 1-\alpha,&\end{aligned} \tag{12}$$

where $Z \sim N(0, 1)$ is a standard normal random variable stochastically independent of the random vector Y. Note that considering the exact $F^\star$-test (8) could be more appropriate for constructing the simultaneous tolerance intervals.

Traditionally, the symmetric tolerance intervals are presented in the form

$$\mathcal{T}^{1-\gamma}_{1-\alpha}(x \mid Y, X) = x'\hat{\beta} \mp S\, K(x \mid 1-\alpha, 1-\gamma, Y, X), \tag{13}$$

where $K(x \mid 1-\alpha, 1-\gamma, Y, X)$ is the tolerance factor which depends on particular values of predictors x, confidence coefficient $1-\alpha$, coverage content $(1-\gamma)$, design matrix X, error standard deviation $S = \sqrt{(Y - X\hat{\beta})'(Y - X\hat{\beta})/(n-k)}$, and $\hat{\beta} = (X'X)^{-1}X'Y$, i.e.

$$\begin{aligned}K(x \mid 1-\alpha, 1-\gamma, Y, X) &=\\ &= \frac{1}{S}\left(x'\hat{\beta} - \inf_{(\beta,\sigma)\in C_{1-\alpha}(Y \mid X)} \left\{x'\beta + u_{\gamma/2}\sigma\right\}\right)\\ &= \frac{1}{S}\left(\sup_{(\beta,\sigma)\in C_{1-\alpha}(Y \mid X)} \left\{x'\beta + u_{1-\gamma/2}\sigma\right\} - x'\hat{\beta}\right).\end{aligned} \tag{14}$$

Numerical Algorithm

Derivation of the tolerance bounds (11) requires numerical optimization for given x, α, γ, the design matrix X, and the observed vector y of Y.

A simplified version of the Matlab code used for calculation of the lower tolerance bound, given in (11), as well as the tolerance factor, (14), is based on the function:

`FMINCON` Minimizes the function `FUN` subject to the nonlinear constraints defined in `NONLCON` (Optimization Toolbox).

```
function [lowerBound, factor] = ...
  LowerBound(x,y,X,quantileU,quantileLR)
x = x(:);
newxq = [x;quantileU];
[n,k] = size(X);
betaML = X \ y;
resid = y - X * betaML;
sigmaML = sqrt(resid' * resid / n);
betaSigma0 = [betaML;sigmaML];
fminconOpt = optimset('MaxFunEvals',...
   500*n, 'Display', 'off');
[betaSigmaArgMin,lowerBound] = ...
   fmincon(@(x) fun(x,newxq),betaSigma0,...
   [],[],[],[],[-Inf;-Inf;1.0e-12],[],...
   @(x) nonlcon(x,y,X,sigmaML,n,...
   quantileLR),fminconOpt);
factor = (x' * betaML - lowerBound) / ...
   (sqrt(n / (n-k)) * sigmaML);
%
function y = fun(x,newxq)
y = newxq' * x;
%
function [c,ceq] = ...
  nonlcon(x,y,X,sigmaML,n,quantileLR)
sig2 = x(3)^2;
resid = y - X * x(1:2);
fun = resid' * resid / sig2 - ...
  n * log(sigmaML^2/sig2) - n - quantileLR;
c = fun; ceq = [];
```

An alternative algorithm for the approximate solution based on Monte Carlo simulations, which is especially useful for linear regression models with $k > 2$, was suggested in [11].

4. Example

For numerical illustration we consider fifteen hypothetical pairs of values (x_i, y_i) selected for the speed-orifice problem, as considered and studied in [3] and [5], the measurement values are given in Table 1.

A simple linear regression model $Y_i = \beta_1 + \beta_2 x_i + \sigma Z_i$, with independent errors $Z_i \sim N(0, 1)$, $i = 1, \ldots, 15$ (i.e. $n = 15$ and $k = 2$), was considered for the analysis of the speed-orifice measurements. The ML estimates $\hat{\beta}$ of the regression parameter $\beta = (\beta_1, \beta_2)'$ and $\hat{\sigma}_{ML}$ of the standard deviation σ are

$$\hat{\beta} = (-19041.86, 17929.64)', \quad \hat{\sigma}_{ML} = 121.50.$$

Moreover, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 1.3531$, $S_x = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 0.0292$.

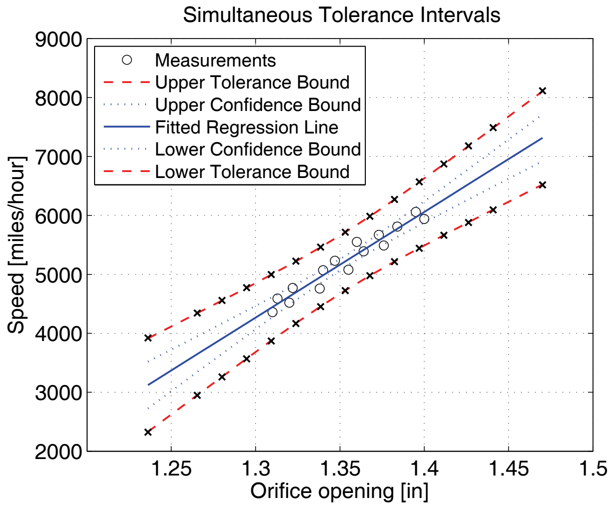

Figure 1. The $(1-\gamma)$-content symmetric simultaneous $(1-\alpha)$-tolerance intervals for the speed-orifice problem calculated at several specified predictors at the significance levels $\alpha = 0.05$ and for prespecified population content $1-\gamma$, with $\gamma = 0.05$. The data and the explicit values of the tolerance factors for the selected simultaneous tolerance bounds, depicted by the symbol $\times$, are presented in Table 1.

For given $k = 2$, $n = 15$, and the significance levels $\alpha = 0.05, 0.01$, the critical values of the test (4) and (8), respectively, for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$, are

$$\lambda_{1-0.05} = 8.6813, \quad \lambda_{1-0.01} = 12.6160,$$
$$F^*_{1-0.05} = 8.1578, \quad F^*_{1-0.01} = 17.3985.$$

Compare with the Table 3 and Table 4. The tables of critical values $F^\star_{1-\alpha}$ are not presented in this paper.

The tolerance factors of the $(1-\gamma)$-content symmetric simultaneous $(1-\alpha)$-tolerance intervals for the speed-orifice problem calculated at several specified predictors f_i, $i = 1, \ldots, 15$, (see the Table 1, columns 3-4, respectively), for different combinations of selected significance levels $\alpha = 0.05$, $\alpha = 0.01$, and population content $1-\gamma$, for $\gamma = 0.05$ and $\gamma = 0.01$, are given in Table 1, (see the columns 5-8).

For illustration, using the function `LowerBound` for the speed-orifice data with the predictor vector $x = (1, 1.3531)$, with the quantile of standard normal distribution $u_{\gamma/2} = -1.96$ for selected significance levels $\gamma = 0.05$ and with the quantile $\lambda_{1-\alpha} = 8.6813$ for selected confidence coefficient $1-\alpha = 0.05$, we get

```
>> [lowerBound,factor] =  ...
   LowerBound([1,1.3531],y,X,-1.96,8.6813)
>> lowerBound = 4.7228e+003
>> factor = 3.7996
```

The upper tolerance bound could be calculated from the tolerance factor, see (13), or directly by using the function `LowerBound` — by minimizing the function $-(x'\beta + u_{1-\gamma/2})$.

Table 1: Tolerance factors of the $(1-\gamma)$-content symmetric simultaneous $(1-\alpha)$-tolerance intervals for the speed-orifice problem calculated at several specified predictors for different combinations of selected significance levels $\alpha = 0.05$, $\alpha = 0.01$, and population content $1-\gamma$, for $\gamma = 0.05$ and $\gamma = 0.01$.

Speed [miles/hour]	Orifice Opening [in]	Predictors [in]	Standardized Predictors	Tolerance Factors $\alpha = 0.05$		Tolerance Factors $\alpha = 0.01$	
y_i	x_i	f_i	$(f_i - \bar{x})/S_x$	$\gamma = 0.05$	$\gamma = 0.01$	$\gamma = 0.05$	$\gamma = 0.01$
4360	1.3100	1.2362	-4	6.1212	7.0053	7.5563	8.5817
4590	1.3130	1.2654	-3	5.3466	6.2590	6.5510	7.6125
4520	1.3200	1.2800	-2.5	4.9779	5.9090	6.0722	7.1578
4770	1.3220	1.2947	-2	4.6298	5.5836	5.6201	6.7348
4760	1.3380	1.3093	-1.5	4.3139	5.2946	5.2095	6.3585
5070	1.3400	1.3239	-1	4.0495	5.0593	4.8654	6.0519
5230	1.3470	1.3385	-0.5	3.8664	4.9014	4.6268	5.8459
5080	1.3550	1.3531	0	3.7996	4.8451	4.5396	5.7723
5550	1.3600	1.3678	0.5	3.8664	4.9014	4.6268	5.8459
5390	1.3640	1.3824	1	4.0495	5.0593	4.8654	6.0519
5670	1.3730	1.3970	1.5	4.3139	5.2946	5.2095	6.3585
5490	1.3760	1.4116	2	4.6298	5.5836	5.6201	6.7348
5810	1.3840	1.4262	2.5	4.9779	5.9090	6.0722	7.1578
6060	1.3950	1.4408	3	5.3466	6.2590	6.5510	7.6125
5940	1.4000	1.4701	4	6.1212	7.0053	7.5563	8.5817

5. Discussion

The presented exact likelihood ratio test (4) for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$ on the parameters β and σ of the linear regression model $Y = X\beta + \sigma Z$ with normally distributed errors, $Z \sim N(0, I_n)$, is especially useful for construction of the simultaneous confidence region for the regression parameters, see (10), and for the simultaneous tolerance intervals (11), and/or the tolerance factors (14), respectively, for future observations predicted for any value of the predictor vector x.

If compared with the $(1-\gamma)$-content symmetric simultaneous $(1-\alpha)$-tolerance intervals based on the product-set approach, as proposed by Limam and Thomas in [5], the suggested tolerance intervals given by (11) are slightly broader for the predictor values close to the middle of the fitted regression function, and became narrower for more distant predictors.

Acknowledgments

The research was supported by the grants VEGA 1/0077/09, VEGA 2/7087/27, APVV-SK-AT-0003-09, and APVV-RPEU-0008-06.

References

[1] De Gryze, S., Langhans, I. & Vandebroek, M. (2007). Using the correct intervals for prediction: A tutorial on tolerance intervals for ordinary least-squares regression. *Chemometrics and Intelligent Laboratory Systems* 87, 147–154.

[2] Krishnamoorthy, K. & Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications, and Computation*. Wiley.

[3] Lieberman, G. J. & Miller, R. G. Jr. (1963). Simultaneous tolerance intervals in regression. *Biometrika* 50, 155–168.

[4] Lieberman, G. J., Miller, R. G. Jr. & Hamilton, A. (1967). Unlimited simultaneous discrimination intervals in regression. *Biometrika* 54, 133–145. Corrections in Biometrika 58, 687.

[5] Limam, M. M. T. & Thomas, D. R. (1988). Simultaneous tolerance intervals for the linear regression model. *Journal of the American Statistical Association* 83 (403), 801–804.

[6] Mee R. W., Eberhardt, K. R. & Reeve, C. P. (1991). Calibration and simultaneous tolerance intervals for regression. *Technometrics* 33 (2), 211–219.

[7] Scheffé, H. (1973). A statistical theory of calibration. *Annals of Statistics* 1(1), 1–37.

[8] Wallis, W. A. (1951). Tolerance intervals for linear regression. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley CA: University of California Press, Berkeley.

[9] Wilks, S. S. (1942). Statistical prediction with special reference to the problem of tolerance limits. *The Annals of Mathematical Statistics* 13, 400–409.

[10] Wilson, A. L. (1967). An approach to simultaneous tolerance intervals in regression. *The Annals of Mathematical Statistics* 38(5), 1536–1540.

[11] Witkovský, V. & Chvosteková M. (2009). Simultaneous tolerance intervals for the linear regression model. In *MEASUREMENT 2009. Proceedings of the International Conference on Measurement*, Smolenice, May 20-23, 2009.

Table 2: Critical values of the likelihood ratio test (LRT) for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$ on parameters of the normal linear regression model with k, $k = 1, \ldots, 10$, explanatory variables, selected small sample sizes n, $n = k + 1, \ldots, 100$, and the significance level $\alpha = 0.1$.

n/k	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2	9.0856	-	-	-	-	-	-	-	-	-
3	6.8029	15.1840	-	-	-	-	-	-	-	-
4	6.0549	10.6953	21.6002	-	-	-	-	-	-	-
5	5.6856	9.2154	14.7996	28.2608	-	-	-	-	-	-
6	5.4659	8.4775	12.5243	19.0877	35.1190	-	-	-	-	-
7	5.3203	8.0346	11.3749	15.9760	23.5341	42.1428	-	-	-	-
8	5.2168	7.7393	10.6781	14.3856	19.5577	28.1187	49.3085	-	-	-
9	5.1395	7.5279	10.2095	13.4126	17.5048	23.2566	32.8254	56.5988	-	-
10	5.0795	7.3695	9.8722	12.7532	16.2386	20.7249	27.0617	37.6412	64.0000	-
11	5.0316	7.2460	9.6176	12.2757	15.3748	19.1522	24.0382	30.9634	42.5556	71.5009
12	4.9925	7.1471	9.4185	11.9136	14.7461	18.0733	22.1485	27.4378	34.9535	47.5599
13	4.9599	7.0662	9.2585	11.6292	14.2670	17.2841	20.8455	25.2223	30.9172	39.0252
14	4.9324	6.9986	9.1271	11.3998	13.8893	16.6803	19.8882	23.6877	28.3688	34.4711
15	4.9089	6.9414	9.0171	11.2108	13.5838	16.2027	19.1532	22.5561	26.5963	31.5834
16	4.8885	6.8924	8.9239	11.0524	13.3313	15.8151	18.5701	21.6843	25.2847	29.5677
17	4.8707	6.8499	8.8437	10.9176	13.1190	15.4940	18.0955	20.9906	24.2713	28.0715
18	4.8550	6.8126	8.7740	10.8015	12.9381	15.2236	17.7014	20.4246	23.4629	26.9123
19	4.8410	6.7798	8.7129	10.7005	12.7820	14.9925	17.3687	19.9537	22.8018	25.9853
20	4.8286	6.7505	8.6589	10.6117	12.6459	14.7928	17.0839	19.5553	22.2505	25.2257
21	4.8173	6.7244	8.6108	10.5332	12.5262	14.6183	16.8374	19.2137	21.7834	24.5911
22	4.8072	6.7008	8.5678	10.4631	12.4200	14.4647	16.6218	18.9176	21.3823	24.0524
23	4.7979	6.6795	8.5289	10.4002	12.3253	14.3283	16.4315	18.6581	21.0339	23.5892
24	4.7895	6.6601	8.4937	10.3435	12.2401	14.2063	16.2625	18.4290	20.7283	23.1863
25	4.7818	6.6424	8.4617	10.2921	12.1633	14.0967	16.1112	18.2251	20.4581	22.8325
26	4.7747	6.6262	8.4324	10.2453	12.0935	13.9976	15.9750	18.0424	20.2174	22.5193
27	4.7682	6.6112	8.4056	10.2024	12.0298	13.9075	15.8517	17.8779	20.0016	22.2400
28	4.7621	6.5974	8.3808	10.1630	11.9716	13.8253	15.7396	17.7288	19.8069	21.9893
29	4.7565	6.5847	8.3580	10.1268	11.9180	13.7500	15.6373	17.5930	19.6304	21.7630
30	4.7513	6.5728	8.3368	10.0932	11.8686	13.6807	15.5434	17.4690	19.4696	21.5576
31	4.7464	6.5617	8.3171	10.0621	11.8229	13.6168	15.4569	17.3552	19.3225	21.3704
32	4.7418	6.5514	8.2987	10.0332	11.7805	13.5576	15.3771	17.2503	19.1874	21.1990
33	4.7375	6.5418	8.2816	10.0062	11.7411	13.5026	15.3032	17.1534	19.0629	21.0415
34	4.7335	6.5327	8.2656	9.9810	11.7043	13.4514	15.2345	17.0636	18.9478	20.8962
35	4.7297	6.5242	8.2505	9.9574	11.6699	13.4037	15.1706	16.9801	18.8411	20.7618
36	4.7261	6.5162	8.2364	9.9353	11.6376	13.3590	15.1108	16.9023	18.7418	20.6371
37	4.7228	6.5087	8.2231	9.9145	11.6074	13.3172	15.0549	16.8296	18.6492	20.5211
38	4.7196	6.5015	8.2105	9.8948	11.5789	13.2778	15.0025	16.7616	18.5627	20.4129
39	4.7166	6.4948	8.1987	9.8763	11.5521	13.2408	14.9532	16.6977	18.4816	20.3117
40	4.7137	6.4884	8.1874	9.8588	11.5267	13.2059	14.9068	16.6377	18.4056	20.2168
45	4.7013	6.4609	8.1392	9.7840	11.4187	13.0576	14.7104	16.3844	18.0857	19.8197
50	4.6915	6.4391	8.1012	9.7252	11.3343	12.9422	14.5583	16.1893	17.8408	19.5174
55	4.6835	6.4214	8.0705	9.6778	11.2664	12.8499	14.4371	16.0345	17.6472	19.2795
60	4.6771	6.4068	8.0451	9.6388	11.2108	12.7743	14.3382	15.9086	17.4903	19.0873
65	4.6712	6.3945	8.0238	9.6061	11.1643	12.7113	14.2559	15.8041	17.3605	18.9288
70	4.6664	6.3840	8.0056	9.5784	11.1248	12.6580	14.1865	15.7161	17.2514	18.7958
75	4.6623	6.3749	7.9900	9.5545	11.0909	12.6123	14.1271	15.6410	17.1584	18.6827
80	4.6587	6.3670	7.9764	9.5337	11.0615	12.5727	14.0756	15.5760	17.0781	18.5852
85	4.6555	6.3600	7.9644	9.5155	11.0357	12.5381	14.0307	15.5193	17.0081	18.5004
90	4.6527	6.3539	7.9538	9.4994	11.0130	12.5075	13.9910	15.4694	16.9466	18.4259
95	4.6502	6.3484	7.9444	9.4850	10.9927	12.4803	13.9558	15.4251	16.8921	18.3599
100	4.6479	6.3434	7.9359	9.4722	10.9746	12.4559	13.9244	15.3855	16.8434	18.3012
∞	4.6052	6.2514	7.7794	9.2364	10.6446	12.0170	13.3616	14.6837	15.9872	17.2750

Table 3: Critical values of the likelihood ratio test (LRT) for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$ on parameters of the normal linear regression model with k, $k = 1, \ldots, 10$, explanatory variables, selected small sample sizes n, $n = k + 1, \ldots, 100$, and the significance level $\alpha = 0.05$.

n/k	1	2	3	4	5	6	7	8	9	10
2	11.8545	-	-	-	-	-	-	-	-	-
3	8.8706	19.3470	-	-	-	-	-	-	-	-
4	7.8893	13.4989	27.1540	-	-	-	-	-	-	-
5	7.4046	11.5844	18.3296	35.2052	-	-	-	-	-	-
6	7.1164	10.6358	15.4138	23.3410	43.4538	-	-	-	-	-
7	6.9257	10.0694	13.9556	19.3806	28.5094	51.8679	-	-	-	-
8	6.7901	9.6927	13.0778	17.3814	23.4748	33.8153	60.4240	-	-	-
9	6.6888	9.4241	12.4904	16.1687	20.9120	27.6847	39.2429	69.1047	-	-
10	6.6103	9.2228	12.0691	15.3518	19.3464	24.5412	31.9998	44.7793	77.8962	-
11	6.5477	9.0663	11.7521	14.7630	18.2858	22.6089	28.2622	36.4110	50.4142	86.7873
12	6.4966	8.9412	11.5047	14.3179	17.5176	21.2931	25.9522	32.0684	40.9101	56.1388
13	6.4541	8.8388	11.3063	13.9693	16.9345	20.3360	24.3719	29.3717	35.9540	45.4906
14	6.4182	8.7536	11.1435	13.6888	16.4763	19.6068	23.2179	27.5192	32.8629	39.9134
15	6.3874	8.6813	11.0075	13.4581	16.1065	19.0319	22.3357	26.1616	30.7317	36.4216
16	6.3608	8.6195	10.8923	13.2649	15.8015	18.5667	21.6383	25.1207	29.1648	34.0061
17	6.3375	8.5659	10.7933	13.1008	15.5456	18.1821	21.0724	24.2955	27.9601	32.2250
18	6.3170	8.5190	10.7074	12.9596	15.3278	17.8587	20.6035	23.6244	27.0028	30.8522
19	6.2988	8.4776	10.6321	12.8369	15.1400	17.5829	20.2085	23.0673	26.2225	29.7589
20	6.2825	8.4408	10.5656	12.7292	14.9765	17.3448	19.8711	22.5971	25.5736	28.8659
21	6.2679	8.4079	10.5064	12.6339	14.8329	17.1371	19.5793	22.1946	25.0249	28.1220
22	6.2546	8.3783	10.4533	12.5490	14.7056	16.9544	19.3244	21.8462	24.5546	27.4919
23	6.2426	8.3515	10.4056	12.4728	14.5920	16.7923	19.0998	21.5414	24.1468	26.9512
24	6.2316	8.3271	10.3623	12.4041	14.4901	16.6475	18.9004	21.2725	23.7897	26.4816
25	6.2216	8.3048	10.3230	12.3419	14.3981	16.5175	18.7221	21.0335	23.4743	26.0700
26	6.2123	8.2844	10.2870	12.2852	14.3146	16.3999	18.5618	20.8196	23.1936	25.7060
27	6.2038	8.2657	10.2540	12.2334	14.2386	16.2932	18.4167	20.6270	22.9421	25.3817
28	6.1959	8.2483	10.2236	12.1858	14.1689	16.1959	18.2849	20.4526	22.7155	25.0910
29	6.1885	8.2323	10.1955	12.1420	14.1050	16.1067	18.1646	20.2941	22.5102	24.8288
30	6.1817	8.2174	10.1695	12.1014	14.0460	16.0247	18.0543	20.1492	22.3234	24.5911
31	6.1753	8.2035	10.1454	12.0639	13.9915	15.9491	17.9528	20.0163	22.1525	24.3745
32	6.1694	8.1906	10.1229	12.0290	13.9409	15.8791	17.8592	19.8940	21.9958	24.1764
33	6.1638	8.1785	10.1019	11.9964	13.8938	15.8142	17.7725	19.7811	21.8513	23.9944
34	6.1586	8.1671	10.0822	11.9660	13.8500	15.7538	17.6919	19.6764	21.7179	23.8267
35	6.1536	8.1564	10.0638	11.9376	13.8089	15.6974	17.6170	19.5792	21.5941	23.6717
36	6.1490	8.1464	10.0464	11.9108	13.7705	15.6447	17.5470	19.4886	21.4791	23.5278
37	6.1446	8.1369	10.0301	11.8857	13.7345	15.5952	17.4815	19.4040	21.3719	23.3941
38	6.1404	8.1280	10.0147	11.8621	13.7005	15.5488	17.4201	19.3248	21.2718	23.2693
39	6.1363	8.1195	10.0002	11.8398	13.6686	15.5052	17.3624	19.2505	21.1780	23.1528
40	6.1328	8.1115	9.9864	11.8187	13.6384	15.4640	17.3081	19.1807	21.0900	23.0435
45	6.1167	8.0770	9.9274	11.7285	13.5098	15.2892	17.0783	18.8863	20.7204	22.5867
50	6.1038	8.0497	9.8809	11.6577	13.4094	15.1533	16.9006	18.6599	20.4377	22.2394
55	6.0934	8.0276	9.8433	11.6007	13.3288	15.0446	16.7591	18.4803	20.2144	21.9663
60	6.0847	8.0092	9.8122	11.5537	13.2627	14.9557	16.6437	18.3343	20.0335	21.7459
65	6.0774	7.9938	9.7861	11.5144	13.2074	14.8816	16.5478	18.2134	19.8840	21.5643
70	6.0712	7.9806	9.7640	11.4811	13.1606	14.8190	16.4668	18.1115	19.7584	21.4120
75	6.0658	7.9693	9.7448	11.4523	13.1204	14.7653	16.3975	18.0244	19.6513	21.2824
80	6.0611	7.9594	9.7282	11.4274	13.0855	14.7187	16.3376	17.9493	19.5590	21.1709
85	6.0570	7.9507	9.7136	11.4055	13.0549	14.6780	16.2852	17.8837	19.4785	21.0738
90	6.0533	7.9429	9.7006	11.3861	13.0279	14.6421	16.2391	17.8259	19.4078	20.9886
95	6.0500	7.9360	9.6891	11.3689	13.0038	14.6102	16.1981	17.7747	19.3451	20.9132
100	6.0470	7.9299	9.6788	11.3534	12.9823	14.5816	16.1615	17.7290	19.2892	20.8460
∞	5.9915	7.8147	9.4877	11.0705	12.5916	14.0671	15.5073	16.9190	18.3070	19.6751

Table 4: Critical values of the likelihood ratio test (LRT) for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$ on parameters of the normal linear regression model with k, $k = 1, \ldots, 10$, explanatory variables, selected small sample sizes n, $n = k + 1, \ldots, 100$, and the significance level $\alpha = 0.01$.

n/k	1	2	3	4	5	6	7	8	9	10
2	18.2902	-	-	-	-	-	-	-	-	-
3	13.6860	29.0049	-	-	-	-	-	-	-	-
4	12.1608	19.9602	40.0323	-	-	-	-	-	-	-
5	11.4053	17.0089	26.4262	51.3036	-	-	-	-	-	-
6	10.9553	15.5571	21.9730	33.0704	62.7725	-	-	-	-	-
7	10.6585	14.6967	19.7741	27.0678	39.8704	74.4068	-	-	-	-
8	10.4474	14.1276	18.4648	24.0871	32.2873	46.8075	86.1831	-	-	-
9	10.2898	13.7236	17.5960	22.3031	28.5000	37.6208	53.8658	98.0839	-	-
10	10.1676	13.4219	16.9772	21.1139	26.2217	33.0087	43.0585	61.0328	110.0954	-
11	10.0705	13.1882	16.5139	20.2636	24.6963	30.2212	37.6075	48.5916	68.2983	122.2068
12	9.9911	13.0017	16.1540	19.6250	23.6016	28.3473	34.2989	42.2904	54.2123	75.6532
13	9.9249	12.8495	15.8663	19.1275	22.7767	26.9976	32.0664	38.4511	47.0516	59.9139
14	9.8693	12.7229	15.6308	18.7288	22.1322	25.9775	30.4533	35.8518	42.6737	51.8861
15	9.8217	12.6160	15.4351	18.4020	21.6145	25.1784	29.2306	33.9680	39.7007	46.9629
16	9.7804	12.5245	15.2692	18.1293	21.1893	24.5350	28.2704	32.5365	37.5402	43.6102
17	9.7443	12.4452	15.1271	17.8982	20.8337	24.0055	27.4956	31.4097	35.8945	41.1679
18	9.7126	12.3760	15.0040	17.6998	20.5319	23.5619	26.8567	30.4986	34.5963	39.3033
19	9.6845	12.3149	14.8962	17.5276	20.2725	23.1847	26.3205	29.7459	33.5447	37.8296
20	9.6593	12.2606	14.8011	17.3768	20.0470	22.8600	25.8640	29.1133	32.6744	36.6337
21	9.6366	12.2122	14.7165	17.2436	19.8493	22.5775	25.4704	28.5738	31.9419	35.6425
22	9.6162	12.1686	14.6409	17.1251	19.6744	22.3294	25.1275	28.1081	31.3163	34.8069
23	9.5975	12.1286	14.5728	17.0189	19.5186	22.1097	24.8260	27.7018	30.7755	34.0924
24	9.5805	12.0933	14.5112	16.9233	19.3790	21.9139	24.5588	27.3443	30.3034	33.4741
25	9.5655	12.0606	14.4552	16.8368	19.2531	21.7382	24.3204	27.0271	29.8874	32.9336
26	9.5506	12.0306	14.4040	16.7579	19.1390	21.5796	24.1063	26.7438	29.5180	32.4569
27	9.5376	12.0031	14.3571	16.6859	19.0352	21.4358	23.9129	26.4892	29.1877	32.0333
28	9.5253	11.9773	14.3140	16.6199	18.9402	21.3048	23.7374	26.2590	28.8906	31.6542
29	9.5140	11.9540	14.2742	16.5591	18.8530	21.1850	23.5773	26.0499	28.6219	31.3129
30	9.5033	11.9323	14.2373	16.5029	18.7728	21.0749	23.4308	25.8592	28.3776	31.0040
31	9.4936	11.9119	14.2031	16.4509	18.6985	20.9734	23.2962	25.6845	28.1546	30.7230
32	9.4845	11.8927	14.1712	16.4026	18.6297	20.8796	23.1720	25.5238	27.9502	30.4664
33	9.4758	11.8753	14.1414	16.3576	18.5658	20.7926	23.0572	25.3755	27.7621	30.2309
34	9.4677	11.8586	14.1136	16.3155	18.5062	20.7116	22.9506	25.2383	27.5885	30.0142
35	9.4601	11.8430	14.0875	16.2762	18.4506	20.6362	22.8514	25.1109	27.4277	29.8140
36	9.4528	11.8283	14.0629	16.2393	18.3984	20.5657	22.7589	24.9923	27.2784	29.6285
37	9.4462	11.8145	14.0398	16.2046	18.3495	20.4996	22.6724	24.8817	27.1393	29.4561
38	9.4393	11.8013	14.0181	16.1719	18.3036	20.4376	22.5914	24.7782	27.0094	29.2956
39	9.4337	11.7889	13.9976	16.1411	18.2603	20.3793	22.5153	24.6811	26.8879	29.1456
40	9.4279	11.7773	13.9779	16.1120	18.2194	20.3244	22.4437	24.5900	26.7740	29.0052
45	9.4031	11.7269	13.8946	15.9877	18.0455	20.0914	22.1411	24.2063	26.2963	28.4191
50	9.3833	11.6869	13.8290	15.8903	17.9098	19.9106	21.9075	23.9117	25.9318	27.9747
55	9.3673	11.6546	13.7760	15.8118	17.8013	19.7662	21.7218	23.6785	25.6444	27.6260
60	9.3538	11.6279	13.7322	15.7474	17.7119	19.6482	21.5706	23.4893	25.4120	27.3450
65	9.3427	11.6054	13.6955	15.6935	17.6375	19.5500	21.4450	23.3326	25.2202	27.1137
70	9.3330	11.5862	13.6643	15.6476	17.5745	19.4670	21.3391	23.2007	25.0591	26.9200
75	9.3247	11.5697	13.6374	15.6082	17.5204	19.3958	21.2485	23.0882	24.9220	26.7554
80	9.3175	11.5552	13.6140	15.5740	17.4735	19.3343	21.1702	22.9911	24.8038	26.6138
85	9.3109	11.5424	13.5935	15.5440	17.4324	19.2804	21.1019	22.9064	24.7009	26.4907
90	9.3054	11.5312	13.5753	15.5175	17.3961	19.2329	21.0416	22.8318	24.6104	26.3827
95	9.3003	11.5213	13.5590	15.4939	17.3639	19.1907	20.9882	22.7658	24.5304	26.2871
100	9.2958	11.5122	13.5445	15.4727	17.3350	19.1529	20.9404	22.7068	24.4589	26.2020
∞	9.2103	11.3449	13.2767	15.0863	16.8119	18.4753	20.0902	21.6660	23.2093	24.7250