Journal of Modern Applied Statistical Methods

Volume 10 | Issue 1 Article 17

5-1-2011

Inference in Simple Regression for the Intercept Utilizing Prior Information on the Slope

Ayman Baklizi
Qatar University, a.baklizi@qu.edu.qa

Adil E. Yousif
Qatar University, Qatar, aeyousif@qu.edu.qa

Journal of Modern Applied Statistical Methods
May 2011, Vol. 10, No. 1, 192-197

1538 – 9472/11/$95.00

Inference in Simple Regression for the Intercept Utilizing Prior Information on the Slope

Ayman Baklizi Adil E. Yousif
Qatar University, Qatar

Shrinkage type estimators are developed for the intercept parameter of a simple linear regression model and the case when it is suspected a priori that the slope parameter is equal to some specific value is considered. Three different estimators of the intercept parameters are examined. The relative performances of the estimators are investigated based on a simulation study of the biases and mean squared errors. The associated bootstrap confidence intervals are also studied and their performance is evaluated.

Key words: Prior information, preliminary test estimator, shrinkage estimators, p-value, bootstrap intervals.

Introduction

Consider the simple linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ i = 1,\ldots,n$$

where the error terms ε_i are independent and identically distributed as $N(0,\sigma^2)$. The aim is to estimate the intercept parameter β_0 when prior information that the slope parameter, β_1, is equal to some specific value, β_1^0 is uncertain. In the absence of any prior information, the maximum likelihood (equivalent, least squares) estimator of the regression parameters are given by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

and

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

Ayman Baklizi is a Professor in the Department of Mathematics in the Statistics and Physics College. Email: a.baklizi@qu.edu.qa. Adil E. Yousif is an Associate Professor in the Department of Mathematics in the Statistics and Physics College. Email: aeyousif@qu.edu.qa.

If the slope parameter is known and is equal to β_1^0, then the estimator of the intercept parameter β_0 is $\hat{\beta}_0' = \bar{y} - \beta_1^0 \bar{x}$. The additional uncertain prior information about the slope parameter is utilized with a view to produce improved estimators. Khan and Saleh (1997) developed the preliminary test estimator and certain types of shrinkage estimators that utilize a so-called distrust factor and adopted this approach to estimate the slope parameters of two suspected parallel regression models. Bhoj and Ahsanullah (1993, 1994) considered estimating the conditional mean for simple regression model, and Khan and Saleh (1997) discussed the problem of shrinkage preliminary test estimation for the multivariate Student-t regression model.

The Estimators

The idea of preliminary test estimation is based on utilizing the result of a preliminary test in choosing between two alternative estimators. In the case considered herein, the two estimators are the unrestricted estimator $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ and the restricted estimator $\hat{\beta}_0' = \bar{y} - \beta_1^0 \bar{x}$. The preliminary test is for the hypothesis:

$$H_0 : \beta_1 = \beta_1^0 \text{ vs. } H_1 : \beta_1 \neq \beta_1^0.$$

The null hypothesis is rejected for large values $|T|$ where

$$T = \frac{S_{xx}^{1/2}\left(\hat{\beta}_1 - \beta_1^0\right)}{S_n},$$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i,$$

$$S_n^2 = \frac{1}{n-2}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$$

and

$$S_{xx} = \sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2.$$

The preliminary test estimator PTE is defined as

$$\hat{\beta}_0^{PT} = \hat{\beta}_0 I\left(|T| > t_{\alpha/2,n-2}\right) + \hat{\beta}_0' I\left(|T| \le t_{\alpha/2,n-2}\right)$$

where I(A) denotes an indicator function of the set A. The PTE is a convex combination of $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ and $\hat{\beta}_0' = \bar{y} - \beta_1^0\bar{x}$ and depends on the random coefficient $I\left(|T| \le t_{\alpha/2,n-2}\right)$ whose value is 1 when the null hypothesis is not rejected and is 0 otherwise. Thus, the PTE is an extreme compromise between $\hat{\beta}_0$ and $\hat{\beta}_0'$. Moreover, the PTE does not allow a smooth transition between the two extremes. A possible remedy for this is to use an estimator with a continuous weight function. This function could be the P-value of the preliminary test.

The use of the P-value as a continuous weight function in preliminary test estimation was utilized by Baklizi (2004) and Baklizi and Abu-Dayyeh, W. (2003). If $v = P - value = \Pr\left(|T| > t_{\alpha/2,n-2}\right)$, then a shrinkage estimator can be found as follows:

$$\hat{\beta}_0^{PV} = (1 - v)\hat{\beta}_0 + v\hat{\beta}_0'.$$

Another possibility is given by Khan and Saleh (1997) who suggested the estimator

$$\hat{\beta}_0^{SH} = \hat{\beta}_0 + c^*\, \hat{\beta}_1 \bar{x} S_n \Big/ \sqrt{s_{xx}}\left|\hat{\beta}_1\right|$$

where c^* is the value that minimizes the mean squared error of $\hat{\beta}_0^{SH}$. This value is given by

$$c^* = \left(\frac{4}{\pi(n-2)}\right)^{1/2} \frac{\Gamma((n-1)/2)}{\Gamma((n-2)/2)}$$

and does not depend on the significance level of the preliminary test.

The Confidence Intervals

Bootstrap intervals are computer intensive methods based on re-sampling with replacement from original data. Bootstrapping regression models can be constructed and run in several ways. The procedure adopted in this study was to resample with replacement from the pairs $(x_i, y_i), i = 1, \ldots, n.$ Several bootstrap based intervals are discussed in the literature.; the most common are the bootstrap-*t* interval, the percentile interval and the bias corrected and accelerated (BCa) interval.

Let $\tilde{\beta}_0$ be an estimator of β_0 and let $\tilde{\beta}_0^*$ be the estimator calculated from the bootstrap sample. Let z_α^* be the α quantile of the bootstrap distribution of $Z^* = (\tilde{\beta}_0^* - \tilde{\beta}_0)/\hat{\eta}^*$, where $\hat{\eta}^*$ is the estimated standard deviation of $\tilde{\beta}_0$ calculated from the bootstrap sample. The bootstrap-*t* interval for β_0 is given by $\left(\tilde{\beta}_0 - z_{1-\alpha/2}^*\hat{\eta}, \tilde{\beta}_0 - z_{\alpha/2}^*\hat{\eta}\right)$ where z_α^* is determined by simulation.

When a variance estimate of the estimator under consideration is unavailable or difficult to obtain, a modification of the bootstrap-*t* interval is needed. Such a modification is based on using a further bootstrap sample from the original bootstrap sample to estimate the variance or the standard deviation $sd^*\left(\tilde{\beta}_0^*\right)$ of $\tilde{\beta}_0^*$. The modified bootstrap-*t* interval is thus given by;

$$\left(\tilde{\beta}_0 - z_{1-\alpha/2}^* sd^*\left(\tilde{\beta}_0^*\right), \tilde{\beta}_0 - \varepsilon_{\alpha/2}^* sd^*\left(\tilde{\beta}_0^*\right)\right)$$

The percentile interval may be described as follows, let $\tilde{\beta}_0^*$ be an estimate of the intercept parameter calculated from the bootstrap sample. Here the bootstrap distribution of $\tilde{\beta}_0^*$ is simulated by re-sampling repeatedly from the regression model based on the original data and calculating $\tilde{\beta}_{0i}^*, i = 1, \ldots, B$ where B is the number of bootstrap samples. If $\hat{H}$ is the cumulative distribution function of $\tilde{\beta}_0^*$, then the $1-\alpha$ interval is given by

$$\left(\hat{H}^{-1}\left(\frac{\alpha}{2}\right),\ \hat{H}^{-1}\left(1-\frac{\alpha}{2}\right) \right).$$

The bias corrected and accelerated interval is also calculated using the percentiles of the bootstrap distribution of $\tilde{\beta}_0$. The percentiles depend on two numbers, $\hat{a}$ and $\hat{z}_0$, called the acceleration and the bias correction. The interval (BCa) is given by

$$\left(\hat{H}^{-1}(\alpha_1),\ \hat{H}^{-1}(\alpha_2) \right)$$

where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{\alpha/2}}{1-\hat{a}\left(\hat{z}_0 + z_{\alpha/2}\right)} \right),$$

and

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{1-\alpha/2}}{1-\hat{a}\left(\hat{z}_0 + z_{1-\alpha/2}\right)} \right),$$

$\Phi(.)$ is the standard normal cumulative distribution function, z_α is the α quantile of the standard normal distribution. The values of $\hat{a}$ and $\hat{z}_0$ are calculated as follows:

$$\hat{a} = \frac{\sum_{i=1}^{n} \left(\tilde{\beta}_0(.) - \tilde{\beta}_0(i) \right)^3}{6\left\{ \sum_{i=1}^{n} \left(\tilde{\beta}_0(.) - \tilde{\beta}_0(i) \right)^2 \right\}^{3/2}}$$

where $\tilde{\beta}_0(i)$ is the intercept estimator using the original data excluding the i[th] pair and

$$\tilde{\beta}_0(.) = \frac{\sum_{i=1}^{n} \tilde{\beta}_0(i)}{n}.$$

The value of $\hat{z}_0$ is given by

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\left\{ \tilde{\beta}_0^{\ *} < \tilde{\beta}_0 \right\}}{B} \right).$$

Methodology

Simulation Study

A simulation study was designed to evaluate the performance of the shrinkage estimators in terms of their biases and mean squared errors. Results for the preliminary test estimator are included for comparison purposes and the performance of the bootstrap intervals associated with the shrinkage estimators is also studied. The simulations used the sample sizes $n = 15$, 30 and 45. The slope parameter true value was chosen to be $\beta_1 = 0$, 1, 2, 3 and 4, the true value of the intercept parameter was set at $\beta_0 = 0$, and the guess value of the slope is set equal to zero in all cases. The predictor values are generated from the uniform distribution while the error terms are generated from $N(0, \sigma^2)$ with $\sigma^2 = 1$ or 4.

For each combination of the simulation indices 1,000 pairs of (x, y) values were generated and the estimators were calculated. The level of the preliminary test is set to $\alpha = 0.05$. The biases and mean squared errors are calculated as:

$$bias\left(\tilde{\beta}_0\right) = \frac{1}{1000} \sum_{i=1}^{1000} \left(\tilde{\beta}_{0i} - \beta_0 \right)$$

and

$$MSE\left(\tilde{\beta}_0\right) = \frac{1}{1000} \sum_{i=1}^{1000} \left(\tilde{\beta}_{0i} - \beta_0 \right)^2,$$

where β_0 is the true value of the intercept parameter and $\tilde{\beta}_0$ is the shrinkage estimator under consideration.

The performance of the intervals is evaluated in terms of their coverage probabilities

(CP) and expected lengths (EL), which are calculated as follows: For the confidence interval CI,

$$CP = \frac{1}{1000}\sum_{i=1}^{1000} I(\beta_0 \in (LB, UB))$$

and

$$EL = \frac{1}{1000}\sum_{i=1}^{1000}(UB - LB)$$

where LB and UB are the lower and upper bounds of the confidence interval. The nominal coverage probability of each interval is taken as 0.95%. The bootstrap calculations used 500 replications, and the second stage used 25 replications to estimate the variances of the estimators.

Results

The results for biases are shown in Table 1. It appears that $\hat{\beta}_0^{PV}$ has the least bias among the shrinkage estimators. The bias of all estimators increases as the initial guess moves further from the true value to a certain point and then decreases again, and the biases of all estimators decreases as the sample size increases. Results for the MSE performance are shown in Table 2; it appears that $\hat{\beta}_0^{PV}$ also has the best overall performance among the shrinkage estimators. For the confidence intervals, it appears that intervals based on $\hat{\beta}_0^{PV}$ perform better than intervals based on $\hat{\beta}_0^{SH}$ in terms of the attainment of coverage probabilities (see Table 3). Results indicate that the BCa intervals and t-int intervals perform better than the PRC intervals among intervals based on $\hat{\beta}_0^{PV}$. Regarding interval widths, it appears that the t-int intervals are the shortest followed closely by the BCa intervals (see Table 4). The PRC intervals are very wide compared to the other intervals based on $\hat{\beta}_0^{PV}$.

Table 1: Biases of the Estimators

β_1^0	$\hat{\beta}_0$	$\hat{\beta}_0^{PT}$	$\hat{\beta}_0^{SH}$	$\hat{\beta}_0^{PV}$
$n = 15$, $\sigma = 1$				
0.0	-0.010	-0.002	-0.007	-0.009
1.0	-0.008	0.314	0.255	0.066
2.0	0.002	0.316	0.354	0.065
3.0	0.012	0.179	0.375	0.042
$n = 15$, $\sigma = 2$				
0.0	-0.008	-0.004	-0.021	-0.013
1.0	0.007	0.374	0.318	0.100
2.0	0.065	0.661	0.555	0.204
3.0	-0.009	0.715	0.639	0.142
$n = 30$, $\sigma = 1$				
0.0	-0.007	0.000	0.004	-0.003
1.0	-0.002	0.230	0.221	0.047
2.0	0.003	0.093	0.258	0.020
3.0	-0.004	0.001	0.252	-0.002
$n = 30$, $\sigma = 2$				
0.0	0.010	-0.004	-0.001	0.007
1.0	-0.019	0.311	0.278	0.062
2.0	-0.007	0.445	0.435	0.089
3.0	0.010	0.369	0.512	0.079

Table 2: MSEs of the Estimators

β_1^0	$\hat{\beta}_0$	$\hat{\beta}_0^{PT}$	$\hat{\beta}_0^{SH}$	$\hat{\beta}_0^{PV}$
$n=15$, $\sigma=1$				
0.0	0.297	0.127	0.154	0.224
1.0	0.296	0.371	0.273	0.281
2.0	0.316	0.695	0.425	0.355
3.0	0.290	0.627	0.438	0.322
$n=15$, $\sigma=2$				
0.0	1.169	0.485	0.606	0.872
1.0	1.200	0.789	0.801	0.979
2.0	1.238	1.467	1.154	1.164
3.0	1.194	2.211	1.406	1.267
$n=30$, $\sigma=1$				
0.0	0.138	0.063	0.073	0.105
1.0	0.142	0.259	0.165	0.152
2.0	0.139	0.269	0.207	0.153
3.0	0.143	0.153	0.208	0.145
$n=30$, $\sigma=2$				
0.0	0.564	0.250	0.300	0.434
1.0	0.566	0.529	0.441	0.504
2.0	0.580	1.049	0.662	0.620
3.0	0.546	1.239	0.786	0.621

Table 3: Coverage Probabilities of the Intervals

β_1^0	$\hat{\beta}_0^{SH}$			$\hat{\beta}_0^{PV}$		
	t-int	BCa	PRC	t-int	BCa	PRC
$n=15$, $\sigma=1$						
0.0	0.744	0.963	0.925	0.922	0.937	0.950
1.0	0.709	0.910	0.856	0.904	0.931	0.876
2.0	0.787	0.842	0.866	0.876	0.925	0.890
3.0	0.855	0.854	0.913	0.910	0.946	0.935
$n=15$, $\sigma=2$						
0.0	0.744	0.957	0.927	0.928	0.941	0.946
1.0	0.705	0.936	0.900	0.910	0.929	0.920
2.0	0.705	0.905	0.862	0.909	0.932	0.885
3.0	0.757	0.875	0.853	0.881	0.919	0.878
$n=30$, $\sigma=1$						
0.0	0.761	0.969	0.933	0.944	0.955	0.944
1.0	0.764	0.892	0.869	0.918	0.949	0.883
2.0	0.863	0.855	0.926	0.925	0.960	0.951
3.0	0.859	0.841	0.922	0.923	0.956	0.928
$n=30$, $\sigma=2$						
0.0	0.733	0.970	0.928	0.937	0.955	0.946
1.0	0.718	0.945	0.893	0.929	0.946	0.903
2.0	0.756	0.874	0.861	0.904	0.933	0.867
3.0	0.838	0.877	0.901	0.916	0.957	0.920

Table 4: Widths of the Bootstrap Intervals

β_1^0	$\hat{\beta}_0^{SH}$			$\hat{\beta}_0^{PV}$		
	t-int	BCa	PRC	t-int	BCa	PRC
$n=15$, $\sigma=1$						
0.0	1.031	1.744	2.680	2.005	1.849	2.591
1.0	1.452	1.780	2.762	2.035	2.075	2.867
2.0	1.884	2.028	2.962	2.117	2.298	3.213
3.0	2.048	2.173	3.068	2.155	2.424	3.236
$n=15$, $\sigma=2$						
0.0	2.070	3.481	5.228	3.979	3.688	5.016
1.0	2.392	3.431	5.399	4.000	3.876	5.347
2.0	2.813	3.560	5.529	4.056	4.123	5.699
3.0	3.461	3.784	5.698	4.115	4.366	6.141
$n=30$, $\sigma=1$						
-	0.734	1.210	1.586	1.364	1.271	1.504
1.0	1.148	1.301	1.733	1.430	1.517	1.853
2.0	1.410	1.478	1.829	1.490	1.678	1.909
3.0	1.455	1.359	1.796	1.474	1.721	1.786
$n=30$, $\sigma=2$						
0.0	1.356	2.429	3.169	2.733	2.526	3.001
1.0	1.832	2.419	3.272	2.772	2.782	3.258
2.0	2.286	2.604	3.441	2.855	3.031	3.654
3.0	2.659	2.838	3.596	2.943	3.236	3.871

Conclusion

In conclusion, it is recommended that $\hat{\beta}_0^{PV}$ and the associated t- interval be employed for inference about the intercept parameter when there uncertain prior information exists regarding the slope.

References

Baklizi, A. (2004). Shrinkage estimation of the common location parameter of several exponentials. *Communications in Statistics, Simulation and Computation*, *33*(*2*), 321-339.

Baklizi, A., & Abu-Dayyeh, W. (2003). Shrinkage estimation of P(X<Y) in the exponential case. *Communications in Statistics, Simulation and Computaion*, *32*(*1*), 31-42.

Bhoj, D. S., & Ahsanullah, M. (1993). Estimation of conditional mean in a linear regression model. *Biometrical Journal*, *35*, 791-799.

Bhoj, D. S., & Ahsanullah, M. (1994). Estimation of a conditional mean in a regression model after a preliminary test on regression coefficient. *Biometrical Journal*, *36*, 153-163.

Khan, S., & Saleh, A. K. Md. E. (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors. *Biometrical Journal*, *29*, 131-147.