STATISTICAL COMPUTING METHODS

Implementaion Mean Regression and Median regression by different approach

Usage
I made this simulation code using `R` . To see regression result, running `main.R` is enough. You can give several options to simulation file `main.R` .

- Number of data [defalut = 1000]
- Value of bias coefficient [defalut = 1]
- Value of beta coefficient [defalut = 1]
- Mode of regression (mean / median) [defalut = mean]
- Your seed of data generation [defalut = 123]

For reporting result, I used following forms of code for mean and median regression each.

```
Rscript main.R -n 1000 -a 1 -b 1 -m mean -s 123
Rscript main.R -n 1000 -a 1 -b 1 -m median -s 123
```

Result reporting

For reporting result, I generated 100 different data sets varying seed inputs(1~100) and measured below things regarding confidence interval:

- Ratio of containing true value
- Distance from true value

Mean Regression

For Mean Regression, I used five methods for comparision.

- Least square method
- Bootstrap
  - Standard method
  - Quantile method
  - Bias-Corrected method
- Bayesian approach using normal likelihood

Used model is $y_i = \alpha + \beta x_i + e_i$, I set $alpha = 1,\ beta = 1,\ e_i \sim N(0,1)$.

I simulated 100 times. And find average of each criteria. In this setting, the perfomances of all methods are similar.

Tables	Containing ratio	Length
LMS_alpha	0.97	0.1236396
LMS_beta	0.94	0.1233002
Boot_standard_alpha	0.97	0.1232451
Boot_standard_beta	0.94	0.1230878
Boot_Quantile_alpha	0.97	0.1229101
Boot_Quantile_beta	0.94	0.1225572
Boot_Bias-Corr_alpha	0.97	0.1229849
Boot_Bias-Corr_beta	0.94	0.1224793
Bayesian	0.97	0.1241528
Bayesian	0.94	0.1211719

Median Regression

For Median Regression, I used four methods for comparision.

- Bootstrap
  - Standard method
  - Quantile method
  - Bias-Corrected method
- Bayesian approach using laplace likelihood

Unlike mean regression case, there are no known conjugate distribution for updating prior multipled by laplace likelihood. Therefore, I should use Metropolis–Hastings algorithm within gibbs sampler. For this, I computed kernel form of each posterior.(σ, α, β) In this process, I set prior distribution to $U(0,1), N(0,1), N(0,1)$ for σ, α, β each. Also I set candidate distribution to normal distribution. In this structure, I can sample sigma, alpha and beta:

Gibbs sampler(arg) :

1. Compute posterior distribution of σ, α, β
2. Sampling with posterior distribution
   - $\sigma \sim p(\sigma|\alpha, \beta)$ with $Metropolis - Hasting\, method$
   - $\alpha \sim p(\alpha|\sigma, \beta)$ with $Metropolis - Hasting\, method$
   - $\beta \sim p(\beta|\sigma, \alpha)$ with $Metropolis - Hasting\, method$
3. until stop condition

Same as Mean regression, I simulated 100 times using different data.

Tables	Containing ratio	Length
Boot_standard_alpha	1.00	0.2724038
Boot_standard_beta	1.00	0.2749494
Boot_Quantile_alpha	0.98	0.2583701
Boot_Quantile_beta	0.93	0.2573865
Boot_Bias-Corr_alpha	0.94	0.2704180
Boot_Bias-Corr_beta	0.95	0.2724501
Bayesian	1.00	6.2374035
Bayesian	1.00	6.2197379

Actually, the Bayesian method had very wide confidence interval. I guess setting of prior and candidate distributions are improper.