STAT GU4261/GR5261 Statistical Methods in Finance

Fall 2018

Homeowork 5 Suggested Solution

If you have questions on the solution, please contact Brian at hl2902@columbia.edu

1. (Problem 2)

   (a) `ind` is the index for which the maximum likelihood is achieved. `ind2` are the indexes of α (in the Box-Cox power transformation) for which the likelihood ratio test with the null hypothesis setting at these α's will not be rejected.

   (b) If `interp = TRUE`, spline interpolation is used to give a smoother plot.

   (c) MLE is 0.36. (A more precise answer is 0.361)

   (d) $[0.32, 0.40]$. (A more precise answer is $[0.3170, 0.4055]$)

   (e) $[0.31, 0.41]$. (A more precise answer is $[0.3030, 0.4195]$)

   Note: a more precise answer answer could be obtained by changing the code:

   ```
   boxcox(male.earnings ~ 1, lambda = seq(0.3, 0.45, 1 / 100)) # code from the book
   ```

   to

   ```
   boxcox(male.earnings ~ 1, lambda = seq(0.3, 0.45, 1 / 2000))
   ```

2. (Problem 3) Estimate of the degree of freedom is 21.6 and estimate of ξ is 1.65.

3. (Problem 4) From Figure 1, we see that the density estimate using a skewed t distribution is similar to the kernel density estimate. I believe that the skewed t-model provides an adequate fit to `male.earnings`.

4. (Problem 5) The density estimate using a skewed t distribution is similar to that using a skewed GED except for region near the peak of the density. It seems that the skewed t estimate is closer to the kernel estimate and hence is better among the two parametric models.

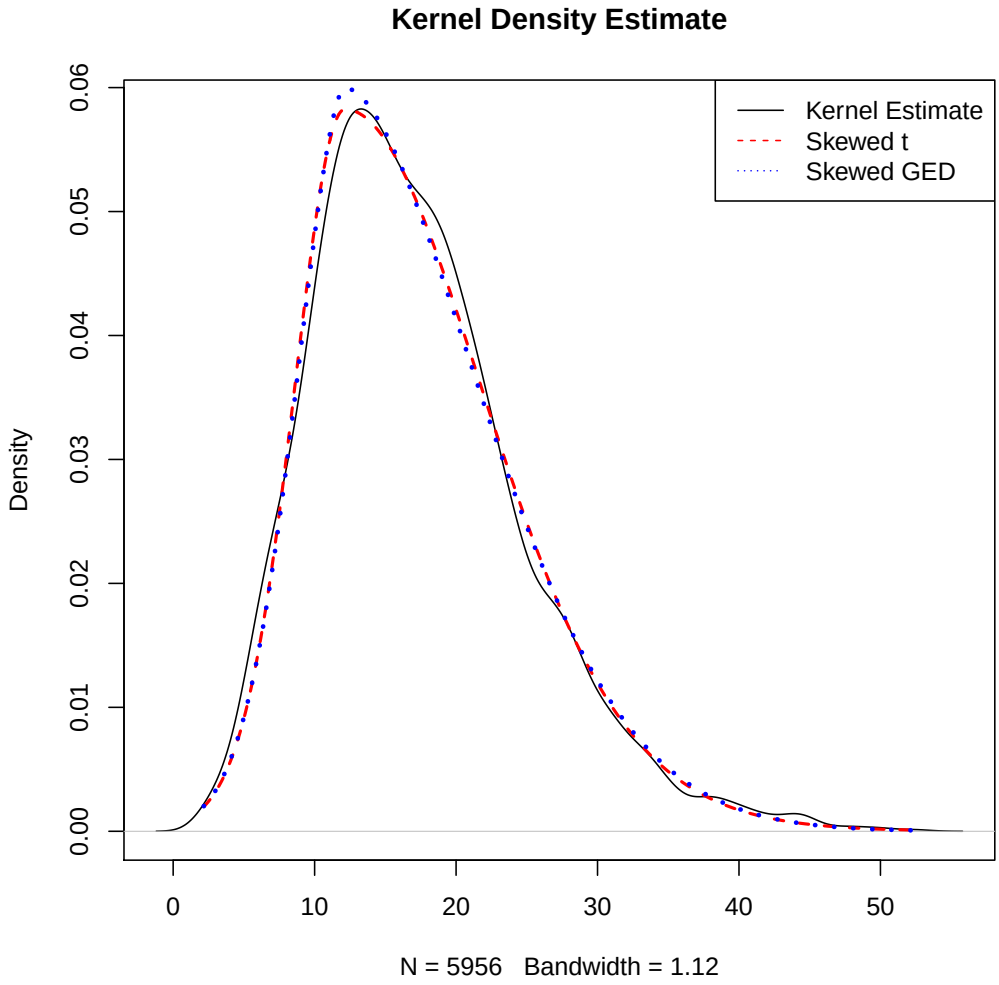
5. (Problem 6) MLE of mean: 0.00078
   MLE of standard deviation: 0.01058
   MLE of degrees-of-freedom parameter 4.03515

   AIC: -11960.47.

6. (Problem 7) MLE of $(\mu, \sigma, \nu, \xi) = (0.00074, 0.01058, 4.03563, 0.98627)$ AIC: -11958.67. AIC selects the t distribution instead of the skewed t-distribution.

   ```
   loglik_sstd = function(x) {
   f = -sum(dsstd(Y, x[1], x[2], x[3], x[4], log = TRUE))
   return(f)
   }
   ```

Figure 1: Problem 4 and Problem 5




```
start = c(mean(Y), sd(Y), 4, 1)
fit_sstd = optim(start, loglik_sstd, method = "L-BFGS-B",
lower = c(-0.1, 0.001, 2.1, 0.001), upper = c(0.1, 1, 20, 10), hessian = TRUE)

minus_logL_std = fit_sstd$value # minus the log-likelihood

AIC_sstd = 2 * minus_logL_std + 2 * length(fit_sstd$par)
```