Tobit v.s. Quantile Analysis

Carver Coleman

April 1, 2020

Regressions and Correlation Matrices

```
tobit_bias_regression <- lm(Coefficient ~ Alpha + Omega + Cutoff, data = tobit)
summary(tobit_bias_regression)
```

```
##
## Call:
## lm(formula = Coefficient ~ Alpha + Omega + Cutoff, data = tobit)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.68108 -0.09257  0.00007  0.09167  1.46015
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.502908   0.003904 128.813   <2e-16 ***
## Alpha       -0.009167   0.000235 -39.008   <2e-16 ***
## Omega        0.000299   0.000235   1.272    0.203
## Cutoff      -0.001002   0.004700  -0.213    0.831
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1946 on 171496 degrees of freedom
## Multiple R-squared:  0.008804,   Adjusted R-squared:  0.008787
## F-statistic: 507.8 on 3 and 171496 DF,  p-value: < 2.2e-16
```

```
tobit_mse_regression <- lm((Coefficient - .5)^2 ~ Alpha + Omega + Cutoff, data = tobit)
summary(tobit_mse_regression)
```

```
##
## Call:
## lm(formula = (Coefficient - 0.5)^2 ~ Alpha + Omega + Cutoff,
##     data = tobit)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09146 -0.03535 -0.00868  0.00997  2.74345
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -8.494e-03  1.525e-03  -5.571 2.53e-08 ***
## Alpha       -1.773e-03  9.177e-05 -19.324  < 2e-16 ***
## Omega        1.511e-02  9.177e-05 164.658  < 2e-16 ***
## Cutoff      -1.715e-02  1.835e-03  -9.342  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.07601 on 171496 degrees of freedom
## Multiple R-squared:  0.1385, Adjusted R-squared:  0.1385
## F-statistic:  9191 on 3 and 171496 DF,  p-value: < 2.2e-16
```

```
quant_bias_regression <- lm(Coefficient ~ Alpha + Omega + Cutoff, data = quantile)
summary(quant_bias_regression)
```

```
##
## Call:
## lm(formula = Coefficient ~ Alpha + Omega + Cutoff, data = quantile)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.82006 -0.11116 -0.00015  0.11223  1.61460
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.923e-01  4.703e-03 104.678   <2e-16 ***
## Alpha       7.718e-05  2.831e-04   0.273    0.785
## Omega       2.594e-04  2.831e-04   0.916    0.360
## Cutoff      7.182e-03  5.662e-03   1.268    0.205
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2345 on 171496 degrees of freedom
## Multiple R-squared:  1.471e-05,  Adjusted R-squared:  -2.783e-06
## F-statistic: 0.8409 on 3 and 171496 DF,  p-value: 0.4712
```

```
quant_mse_regression <- lm((Coefficient - .5)^2 ~ Alpha + Omega + Cutoff, data = quantile)
summary(quant_mse_regression)
```

```
##
## Call:
## lm(formula = (Coefficient - 0.5)^2 ~ Alpha + Omega + Cutoff,
##     data = quantile)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1215 -0.0518 -0.0108  0.0127  3.2151
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.888e-02  2.177e-03 -13.266   <2e-16 ***
## Alpha       -3.761e-05  1.310e-04  -0.287   0.7741
## Omega        2.190e-02  1.310e-04 167.155   <2e-16 ***
## Cutoff      -4.679e-03  2.620e-03  -1.785   0.0742 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1085 on 171496 degrees of freedom
## Multiple R-squared:  0.1401, Adjusted R-squared:  0.1401
## F-statistic:  9315 on 3 and 171496 DF,  p-value: < 2.2e-16
```

```
kable(cor(quantile))
```

	Alpha	Omega	Cutoff	Coefficient
Alpha	1.0000000	0.0000000	0.0000000	0.0006583
Omega	0.0000000	1.0000000	0.0000000	0.0022126
Cutoff	0.0000000	0.0000000	1.0000000	0.0030629
Coefficient	0.0006583	0.0022126	0.0030629	1.0000000

```
kable(cor(tobit))
```

	Alpha	Omega	Cutoff	Coefficient
Alpha	1.0000000	0.0000000	0.0000000	-0.0937795
Omega	0.0000000	1.0000000	0.0000000	0.0030588
Cutoff	0.0000000	0.0000000	1.0000000	-0.0005127
Coefficient	-0.0937795	0.0030588	-0.0005127	1.0000000

Graphics:

Change in Alpha with cutoffs greater than or equal to top 75% of data

```
# Calculate subsetted averages
tobit_alpha <- subset(tobit, tobit$Cutoff >= .75)
tob <- tapply(tobit_alpha$Coefficient, tobit_alpha$Alpha, mean)
quantile_alpha <- subset(quantile, quantile$Cutoff >= .75)
quant <- tapply(quantile_alpha$Coefficient, quantile_alpha$Alpha, mean)

# Create dataframe with bias and MSE
data <- data.frame(-3:3, tob - .5, (tob - .5)^2, "Tobit", row.names = 1:7)
data1 <- data.frame(-3:3, quant - .5, (quant - .5)^2, "Quantile", row.names = 1:nrow(data))
colnames(data) <- c("Alpha", "Bias", "MSE" , "Type")
colnames(data1) <- colnames(data)
final <- as.data.frame(rbind(data, data1))

# Bias graph
jpeg('alpha_bias.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final, aes(x = Alpha, y = Bias, color = Type)) +
  geom_point() +
  geom_path() +
  geom_abline(mapping = aes(slope = 0, intercept = 0)) +
  ylim(-.1, .1) +
  labs(title = "Tobit and Quantile Bias for levels of Skewness",
       subtitle = "Only Cutoffs above 75th percentile")
dev.off()
```

```
## pdf
##   2
```

```r
# MSE graph
jpeg('alpha_mse.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final, aes(x = Alpha, y = MSE, color = Type)) +
  geom_point() +
  geom_path() +
  geom_abline(mapping = aes(slope = 0, intercept = 0)) +
  ylim(-.005, .005) +
  labs(title = "Tobit and Quantile MSE for levels of Skewness",
       subtitle = "Only Cutoffs above 75th percentile")
dev.off()
```

```
## pdf
##   2
```

Change in Omega with 75th Percentile and Above and Base Alpha level

```r
# Calculate subsetted averages
tobit_alpha <- subset(tobit, tobit$Cutoff >= .75 & tobit$Alpha == 0)
tob <- tapply(tobit_alpha$Coefficient, tobit_alpha$Omega, mean)
quantile_alpha <- subset(quantile, quantile$Cutoff >= .75 & quantile$Alpha %in% c(-3,-2,-1))
quant <- tapply(quantile_alpha$Coefficient, quantile_alpha$Omega, mean)

# Create dataframe with bias and MSE
data <- data.frame(1:7, tob - .5, (tob - .5)^2, "Tobit", row.names = 1:nrow(data))
data1 <- data.frame(1:7, quant - .5, (quant - .5)^2, "Quantile", row.names = 1:nrow(data1))
colnames(data) <- c("Omega", "Bias", "MSE", "Type")
colnames(data1) <- colnames(data)
final <- as.data.frame(rbind(data, data1))

# Bias graph
jpeg('omega_bias.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final, aes(x = Omega, y = Bias, color = Type)) +
  geom_point() +
  geom_path() +
  geom_abline(mapping = aes(slope = 0, intercept = 0)) +
  ylim(-.1, .1) +
  labs(title = "Tobit and Quantile Bias for levels of Spread",
       subtitle = "Only Cutoffs above 75th percentile and Base Skewness")
dev.off()
```

```
## pdf
##   2
```

```r
# MSE graph
jpeg('omega_mse.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final, aes(x = Omega, y = MSE, color = Type)) +
  geom_point() +
  geom_path() +
  geom_abline(mapping = aes(slope = 0, intercept = 0)) +
  ylim(-.0005, .0005) +
  labs(title = "Tobit and Quantile MSE for levels of Spread",
       subtitle = "Only Cutoffs above 75th percentile and Base Skewness")
dev.off()
```

```
## pdf
##   2
```

Change in Cutoff with Base Alpha level

```
# Calculate subsetted averages
tobit_alpha <- subset(tobit, tobit$Alpha == 0)
tob <- tapply(tobit_alpha$Coefficient, tobit_alpha$Cutoff, mean)
quantile_alpha <- subset(quantile, quantile$Alpha %in% c(-3,-2,-1))
quant <- tapply(quantile_alpha$Coefficient, quantile_alpha$Cutoff, mean)

# Create dataframe with bias and MSE
data <- data.frame(c(.65, .7, .75, .8, .85, .9, .95), tob - .5, (tob - .5)^2, "Tobit",
                   row.names = 1:nrow(data))
data1 <- data.frame(c(.65, .7, .75, .8, .85, .9, .95), quant - .5, (quant - .5)^2, "Quantile",
                    row.names = 1:nrow(data1))
colnames(data) <- c("Cutoff", "Bias", "MSE", "Type")
colnames(data1) <- colnames(data)
final <- as.data.frame(rbind(data, data1))

# Bias graph
jpeg('cutoff_bias.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final, aes(x = Cutoff, y = Bias, color = Type)) +
  geom_point() +
  geom_path() +
  geom_abline(mapping = aes(slope = 0, intercept = 0)) +
  ylim(-.1, .1) +
  labs(title = "Tobit and Quantile Bias for Percentile Cutoff Levels",
       subtitle = "Only Base Skewness")
dev.off()
```

```
## pdf
##   2
```

```
# MSE graph
jpeg('cutoff_mse.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final, aes(x = Cutoff, y = MSE, color = Type)) +
  geom_point() +
  geom_path() +
  geom_abline(mapping = aes(slope = 0, intercept = 0)) +
  ylim(-.0005, .0005) +
  labs(title = "Tobit and Quantile MSE for Percentile Cutoff Levels",
       subtitle = "Only Base Skewness")
dev.off()
```

```
## pdf
##   2
```

```
quantile <- data.frame(quantile, "Type" = "Quantile")
tobit <- data.frame(tobit, "Type" = "Tobit")
final_data <- as.data.frame(rbind(quantile, tobit))
```

```r
jpeg('density.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
ggplot(final_data, aes(x = Coefficient - .5, color = Type)) +
  geom_density() +
  xlim(-5,5) +
  xlab("Bias") +
  labs(title = "Bias Density of Tobit v.s. Quantile Coefficients")
dev.off()
```

```
## pdf
##   2
```

```r
# quant <- read.csv("quantile_accuracy.csv")
# mean(quant$Difference < .5)
#
# tapply(quant$Difference[quant$Cutoff <= 0.65], quant$Alpha[quant$Cutoff <= .65], mean)
#
# jpeg('density_quant.jpg', quality = 100, width = 12, height = 8, units = "in", res = 300)
# ggplot(quant, aes(x = Difference)) +
#   geom_density() +
#   xlim(0,1) +
#   xlab("Bias") +
#   geom_vline(aes(xintercept = .5, color = "red")) +
#   labs(title = "Bias Density of Tobit v.s. Quantile Coefficients") +
#   theme(legend.position = "none")
# dev.off()
```