A Curveball Index: Quantification of Breaking Balls for Pitchers
Technical Material for Camera Set-up
December 7, 2009

Location: Biola University Production Center

-There are 3 cameras setup to collect data

Camera 1 is placed directly to the side where the pitcher releases the ball.

- Records the initial height of the pitch
-Markers start at 6 Feet. Could start lower for shorter pitchers.
-Markers every 3 inches
-Finishes at 7 feet 6 inches

Camera 2 is placed about 30 feet to the side of the plane of the pitch.

-Records the max height and the breaking point
-Horizontally marked every 4 feet
-Vertical markers every 3 inches
-Vertical markers start at 4 feet and end at 7 feet
-Camera records horizontal markers starting at 8 feet and ending at 32 feet

Camera 3 is place in front of the catcher, facing the catcher.

-Records the final location of the pitch
-Markers every 6 Inches
-From the ground to 6 feet

Post

-Using a capture deck, take the tape and visually pinpoint all the factors.

Raw Data

Pitcher	Intial Height	Max height	Rise (")	Breaking Point (')	Final location	Break (")	Distance from knees (")	Rise Points	Break Points	Distance Points	Location Points	Rating
A	5'9"	6'5"	8	25	3'8"	33	26	320	660	500	520	20
A	5'8"	6'0"	4	23	1'2"	58	4	160	1160	460	80	75
A	5'10"	6'0"	2	18	2'2"	46	8	80	920	360	160	45
A	5'10"	5'10"	0	14.5	1'7"	51	1	0	1020	290	20	56
A	5'9"	5'11"	2	11	0'3"	68	15	80	1360	220	300	35
A	5'9"	5'9.5"	0.5	21	1'	68.5	6	20	1370	420	120	60
A	5'10	5'10"	0	14	0'10"	60	8	0	1200	280	160	57
A	5'11	6'2"	3	21.5	2'2"	48	8	120	960	430	160	52
A	5'9"	6'2"	3	19.5	2'6"	44	12	120	880	390	240	25
A	5'10"	6'7"	9	26.5	4'1"	30	31	360	600	530	620	18
B	6'2"	6'4"	2	17	2'1"	51	7	80	1020	340	140	45
B	6'1"	6'4"	3	17.5	1'10"	66	4	120	1320	350	80	55
B	6'1"	6'8"	7	20.5	3'11"	33	29	280	660	410	580	35
B	6'1"	6'5"	4	24.5	3'4"	37	22	160	740	490	440	48
B	6'0"	7'4"	16	30.5	4'7"	33	37	640	660	610	740	23
B	6'2"	6'3"	1	19.5	2'3"	48	9	40	960	390	180	50
B	6'0"	6'3"	3	23.5	2'9"	42	15	120	840	470	300	56
B	6'0"	6'7"	7	26.5	4'5"	26	35	280	520	530	700	36
B	5'10"	5'10.5"	0.5	15	0"	70.5	18	20	1410	300	360	58
B	6'0"	6'0"	0	17.5	1'3"	57	3	0	1140	350	60	70
C	5'3"	5'4"	1	10	0'4"	60	14	40	1200	200	280	41
C	5'4"	5'4"	0	12	0'7"	57	11	0	1140	240	220	55
C	5'5"	5'5"	0	8	-0'3"	68	21	0	1360	160	420	40
C	5'3"	5'5"	2	20	2'7"	34	13	80	680	400	260	31
C	5'6"	6'6"	12	24	1'10"	56	4	480	1120	480	80	41
C	5'4"	5'5"	1	16.5	1'7"	46	1	40	920	330	20	50

C	5'5"	5'5.5"	0.5	4	-0'3"	68.5	21	20	1370	80	420	32
C	5'4"	5'4"	0	11.5	0'1"	63	17	0	1260	230	340	47
C	5'6"	5'9"	3	25.5	2'1"	44	7	120	880	510	140	65
C	5'4"	5'4	0	23.5	1'10"	42	4	0	840	470	80	70

R Scripts

These scripts may be run in R (www.r-project.org) to reproduce the results of the paper.

```
##### Load Data
# Directly
data <-
  structure(c(20, 75, 45, 56, 35, 60, 57, 52, 25, 18, 45, 55, 35,
         48, 23, 50, 56, 36, 58, 70, 41, 55, 40, 31, 41, 50, 32, 47, 65,
         70, 8, 4, 2, 0, 2, 0.5, 0, 3, 3, 9, 2, 3, 7, 4, 16, 1, 3, 7,
         0.5, 0, 1, 0, 0, 2, 12, 1, 0.5, 0, 3, 0, 33, 58, 46, 51, 68,
         68.5, 60, 48, 44, 30, 51, 66, 33, 37, 33, 48, 42, 26, 70.5, 57,
         60, 57, 68, 34, 56, 46, 68.5, 63, 44, 42, 25, 23, 18, 14.5, 11,
         21, 14, 21.5, 19.5, 26.5, 17, 17.5, 20.5, 24.5, 30.5, 19.5,
23.5,
         26.5, 15, 17.5, 10, 12, 8, 20, 24, 16.5, 4, 11.5, 25.5, 23.5,
         26, 4, 8, 1, 15, 6, 8, 8, 12, 31, 7, 4, 29, 22, 37, 9, 15, 35,
         18, 3, 14, 11, 21, 13, 4, 1, 21, 17, 7, 4), .Dim = c(30L, 5L),
.Dimnames = list(
          NULL, c("rating", "rise", "totalbreak", "breakingpoint",
              "location")))
data = data.frame(data)

#Abbreviations
attach(data)
rs = rise
tb = totalbreak
bkpt = breakingpoint
lc = location

##### Analysis
### Regression Model
cor(cbind(rating, rise,totalbreak,bkpt,location))
pairs(cbind(rating, rise,totalbreak,bkpt,location))
m0 = lm(rating ~ -1 + rs + tb + bkpt + lc)

##### Checking assumptions and collinearity
#Normality of residuals: YES!
model = lm(formula = rating ~ -1 + rs + tb + bkpt + lc)
shapiro.test(model$residuals)
#       Shapiro-Wilk normality test
#
#data:  model$residuals
#W = 0.9705, p-value = 0.5532
library(car)
qqPlot(model$residuals)
title("Normal-QQ Plot of Residuals With Confidence Bands")

#Multicollinearity:
#              rating      rise totalbreak breakingpoint   location
#rating       1.00000000 -0.5437556  0.4058752   -0.06907412 -
0.7032479
#rise        -0.54375562  1.0000000 -0.5583619    0.69210812
0.5703195
#totalbreak    0.40587517 -0.5583619  1.0000000   -0.74369262
-0.4589465
#breakingpoint -0.06907412  0.6921081 -0.7436926
1.00000000  0.2401694
#location     -0.70324787  0.5703195 -0.4589465    0.24016942
1.0000000

kappa(data[,2:5],exact=TRUE)  #24.07946, quite low

# Evaluate homoscedasticity
plot(model$fitted.values, model$residuals)
title("Residuals vs. Fitted Values")
lines(seq(10,70,length=100),rep(0,100),lty="dotted")

# non-constant error variance test: Bruesch-Pagan Test of
Hetereoskedasticity
```

```
library(car)
ncvTest(model)
#Non-constant Variance Score Test
#Variance formula: ~ fitted.values
#Chisquare = 0.02992489    Df = 1     p = 0.8626608

#Graph in the paper:
par(mfrow=c(2,1))
plot(model$fitted.values, model$residuals,xlab="Fitted
values",ylab="Residuals")
title("Residuals vs. Fitted")
lines(seq(10,70,length=100),rep(0,100),lty="dotted")
qqPlot(model$residuals,xlab="Normal Quantiles",
ylab="Residuals")
title("Normal-QQ Plot of Residuals")

# Evaluate Collinearity: All below 10, which is a good sign!
vif(model) # variance inflation factors
#      rs       tb     bkpt       lc
#4.169596 5.617257 9.356591 3.834900
#Warning message:
#In vif.lm(model) : No intercept: vifs may not be sensible.


#####  R^2 for no-intercept model
ybar = mean(rating)
#[1] 46.36667
min(ybar)
#[1] 46.36667
min(rating)
#[1] 18
median(rating)
#[1] 47.5
SST = sum((rating-ybar)^2)
#[1] 6526.967
1-1786/SST
#[1] 0.726366
1786/6527
#[1] 0.2736326


##### CI's
b = c(-2.51, 1.88, -0.47, .51)
qt(0.025,28)
#[1] -2.048407
t = qt(0.025,28)
s = c(0.62227, 0.23530, 0.17489, 0.06905)
cbind(b,b+t*s,b-t*s)/57
#[1,] -0.044035088 -0.066397584 -0.021672591
#[2,]  0.032982456  0.024526488  0.041438425
#[3,] -0.008245614 -0.014530630 -0.001960598
#[4,]  0.008947368  0.006465921  0.011428816
cbind(b,b+t*s,b-t*s)
#[1,] -2.51 -3.7846623 -1.2353377
#[2,]  1.88  1.3980098  2.3619902
#[3,] -0.47 -0.8282459 -0.1117541
#[4,]  0.51  0.3685575  0.6514425
```