

# Africa

Chris Parrish

June 27, 2016

## Contents

<b>Africa</b>	<b>1</b>
data . . . . .	1
map . . . . .	1
include a categorical (indicator) variable in the linear model . . . . .	2
include interaction . . . . .	4
analysis . . . . .	7

Africa

reference: McElreath, Statistical Rethinking, chap 7, pp.209-239

```
library(rethinking)
library(ggplot2)
```

## Africa

### data

```
## R code 7.1
data(rugged)
d <- rugged

# make log version of outcome
d$log_gdp <- log( d$rgdppc_2000 )

# extract countries with GDP data
dd <- d[ complete.cases(d$rgdppc_2000) , ]

# split countries into Africa and not-Africa
d.A1 <- dd[ dd$cont_africa==1 , ] # Africa
d.A0 <- dd[ dd$cont_africa==0 , ] # not Africa
```

### map

```
## R code 7.2
# African nations
m7.1 <- map(
  alist(
    log_gdp ~ dnorm( mu , sigma ) ,
    mu <- a + bR*rugged ,
    a ~ dnorm( 8 , 100 ) ,
    bR ~ dnorm( 0 , 1 ) ,
```

```

    sigma ~ dunif( 0 , 10 )
  ) ,
  data=d.A1 )

```

```

# non-African nations
m7.2 <- map(
  alist(
    log_gdp ~ dnorm( mu , sigma ) ,
    mu <- a + bR*rugged ,
    a ~ dnorm( 8 , 100 ) ,
    bR ~ dnorm( 0 , 1 ) ,
    sigma ~ dunif( 0 , 10 )
  ) ,
  data=d.A0 )

```

All countries

```

## R code 7.3
m7.3 <- map(
  alist(
    log_gdp ~ dnorm( mu , sigma ) ,
    mu <- a + bR*rugged ,
    a ~ dnorm( 8 , 100 ) ,
    bR ~ dnorm( 0 , 1 ) ,
    sigma ~ dunif( 0 , 10 )
  ) ,
  data=dd )

```

All countries, with Africa indicator variable

```

## R code 7.4
m7.4 <- map(
  alist(
    log_gdp ~ dnorm( mu , sigma ) ,
    mu <- a + bR*rugged + bA*cont_africa ,
    a ~ dnorm( 8 , 100 ) ,
    bR ~ dnorm( 0 , 1 ) ,
    bA ~ dnorm( 0 , 1 ) ,
    sigma ~ dunif( 0 , 10 )
  ) ,
  data=dd )

```

## include a categorical (indicator) variable in the linear model

Use a linear model to compare African and non-African countries

```

## R code 7.5
compare( m7.3 , m7.4 )

```

```

##      WAIC pWAIC dWAIC weight   SE  dSE
## m7.4 476.2   4.3   0.0     1 15.32  NA
## m7.3 539.7   2.8  63.5     0 13.31 15.12

```

```

## R code 7.6
rugged.seq <- seq(from=-1,to=8,by=0.25)

```

```
# compute mu over samples, fixing cont_africa=0
mu.NotAfrica <- link( m7.4 , data=data.frame(cont_africa=0,rugged=rugged.seq) )
```

```
## [ 100 / 1000 ]
[ 200 / 1000 ]
[ 300 / 1000 ]
[ 400 / 1000 ]
[ 500 / 1000 ]
[ 600 / 1000 ]
[ 700 / 1000 ]
[ 800 / 1000 ]
[ 900 / 1000 ]
[ 1000 / 1000 ]
```

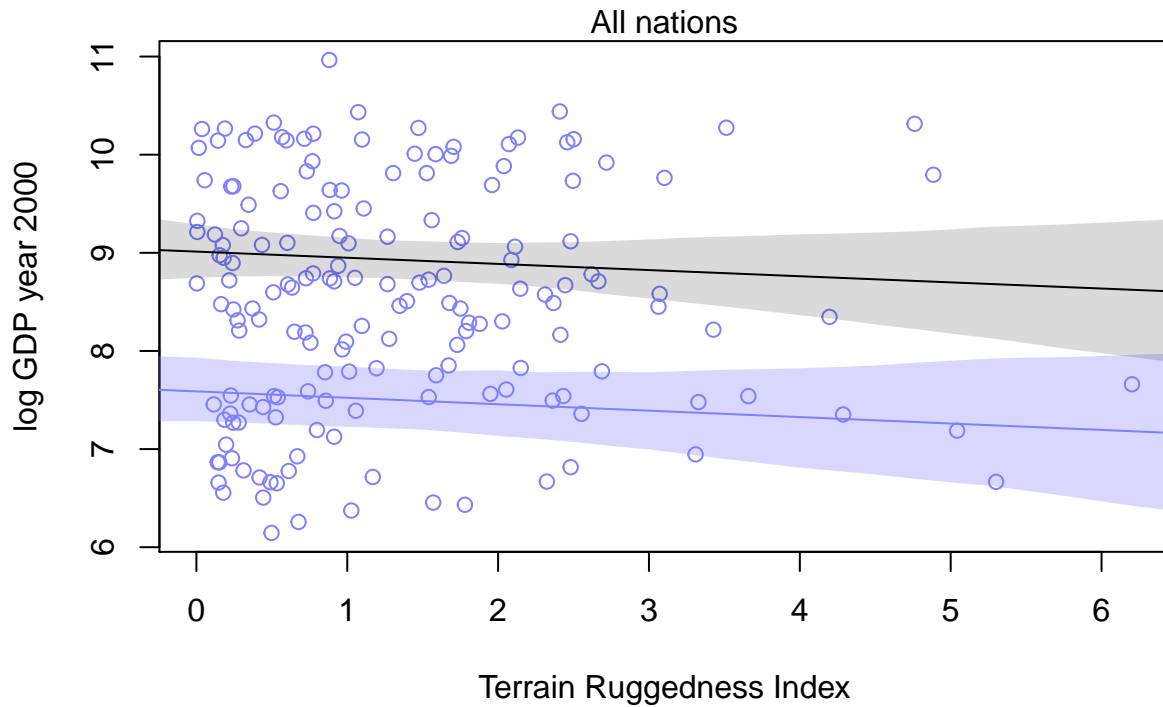
```
# compute mu over samples, fixing cont_africa=1
mu.Africa <- link( m7.4 , data=data.frame(cont_africa=1,rugged=rugged.seq) )
```

```
## [ 100 / 1000 ]
[ 200 / 1000 ]
[ 300 / 1000 ]
[ 400 / 1000 ]
[ 500 / 1000 ]
[ 600 / 1000 ]
[ 700 / 1000 ]
[ 800 / 1000 ]
[ 900 / 1000 ]
[ 1000 / 1000 ]
```

```
# summarize to means and intervals
mu.NotAfrica.mean <- apply( mu.NotAfrica , 2 , mean )
mu.NotAfrica.PI <- apply( mu.NotAfrica , 2 , PI , prob=0.97 )
mu.Africa.mean <- apply( mu.Africa , 2 , mean )
mu.Africa.PI <- apply( mu.Africa , 2 , PI , prob=0.97 )
```

Plot (figure 7.3)

```
plot( log(rgdppc_2000) ~ rugged , data=dd ,
      col=rangi2 , ylab="log GDP year 2000" ,
      xlab="Terrain Ruggedness Index" )
mtext( "All nations" , 3 )
lines( rugged.seq , mu.Africa.mean , col=rangi2 )
shade( mu.Africa.PI , rugged.seq , col=col.alpha(rangi2,0.3) )
lines( rugged.seq , mu.NotAfrica.mean )
shade( mu.NotAfrica.PI , rugged.seq )
```



## include interaction

map with interaction on two lines (expand coefficient into linear model)

```
## R code 7.7
m7.5 <- map(
  alist(
    log_gdp ~ dnorm( mu , sigma ) ,
    mu <- a + gamma*rugged + bA*cont_africa ,
    gamma <- bR + bAR*cont_africa ,
    a ~ dnorm( 8 , 100 ) ,
    bA ~ dnorm( 0 , 1 ) ,
    bR ~ dnorm( 0 , 1 ) ,
    bAR ~ dnorm( 0 , 1 ) ,
    sigma ~ dunif( 0 , 10 )
  ) ,
  data=dd )
```

Compare

```
## R code 7.8
compare( m7.3 , m7.4 , m7.5 )
```

##	WAIC	pWAIC	dWAIC	weight	SE	dSE
## m7.5	469.9	5.4	0.0	0.96	15.18	NA
## m7.4	476.3	4.4	6.4	0.04	15.28	6.24
## m7.3	539.7	2.8	69.8	0.00	13.33	15.15

Interaction on one line (symmetry)

```
## R code 7.9
m7.5b <- map(
  alist(
```

```

log_gdp ~ dnorm( mu , sigma ) ,
mu <- a + bR*rugged + bAR*rugged*cont_africa + bA*cont_africa,
a ~ dnorm( 8 , 100 ) ,
bA ~ dnorm( 0 , 1 ) ,
bR ~ dnorm( 0 , 1 ) ,
bAR ~ dnorm( 0 , 1 ) ,
sigma ~ dunif( 0 , 10 )
) ,
data=dd )

```

Compare

```
## R code 7.10
```

```
rugged.seq <- seq(from=-1,to=8,by=0.25)
```

```
mu.Africa <- link( m7.5 , data=data.frame(cont_africa=1,rugged=rugged.seq) )
```

```
## [ 100 / 1000 ]
[ 200 / 1000 ]
[ 300 / 1000 ]
[ 400 / 1000 ]
[ 500 / 1000 ]
[ 600 / 1000 ]
[ 700 / 1000 ]
[ 800 / 1000 ]
[ 900 / 1000 ]
[ 1000 / 1000 ]
```

```
mu.Africa.mean <- apply( mu.Africa , 2 , mean )
mu.Africa.PI <- apply( mu.Africa , 2 , PI , prob=0.97 )
```

```
mu.NotAfrica <- link( m7.5 , data=data.frame(cont_africa=0,rugged=rugged.seq) )
```

```
## [ 100 / 1000 ]
[ 200 / 1000 ]
[ 300 / 1000 ]
[ 400 / 1000 ]
[ 500 / 1000 ]
[ 600 / 1000 ]
[ 700 / 1000 ]
[ 800 / 1000 ]
[ 900 / 1000 ]
[ 1000 / 1000 ]
```

```
mu.NotAfrica.mean <- apply( mu.NotAfrica , 2 , mean )
mu.NotAfrica.PI <- apply( mu.NotAfrica , 2 , PI , prob=0.97 )
```

```
## R code 7.11
```

```
# plot African nations with regression
```

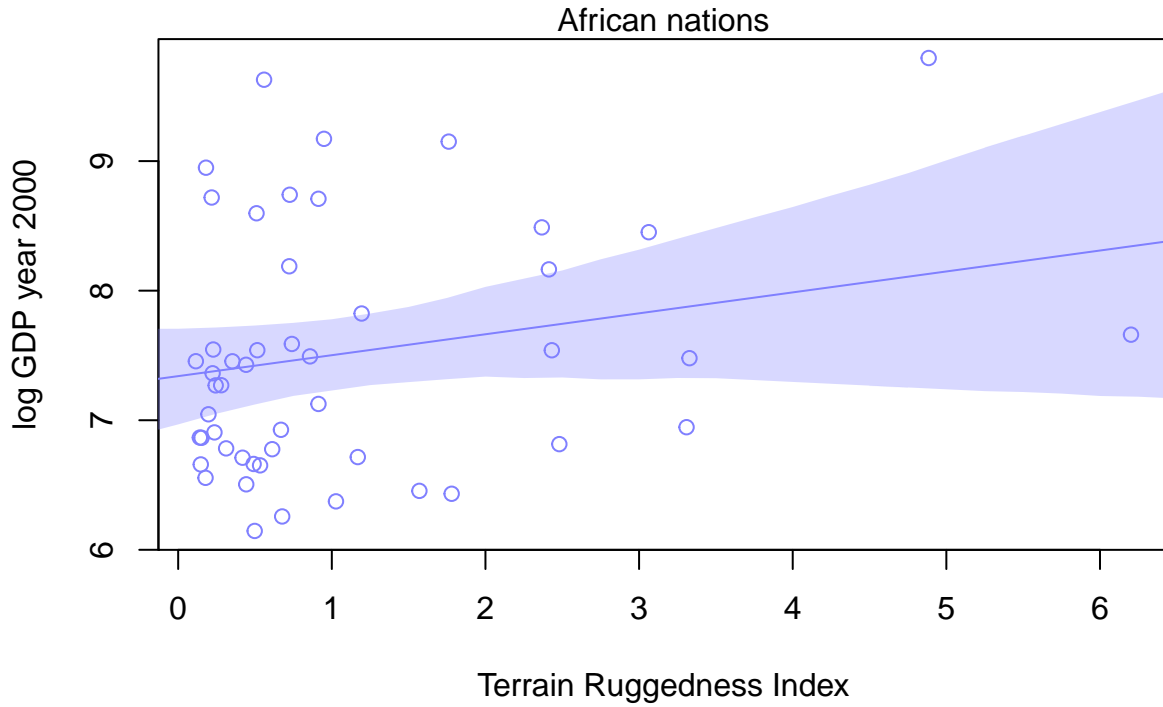
```
d.A1 <- dd[dd$cont_africa==1,]
```

```
plot( log(rgdppc_2000) ~ rugged , data=d.A1 ,
      col=rangi2 , ylab="log GDP year 2000" ,
      xlab="Terrain Ruggedness Index" )
```

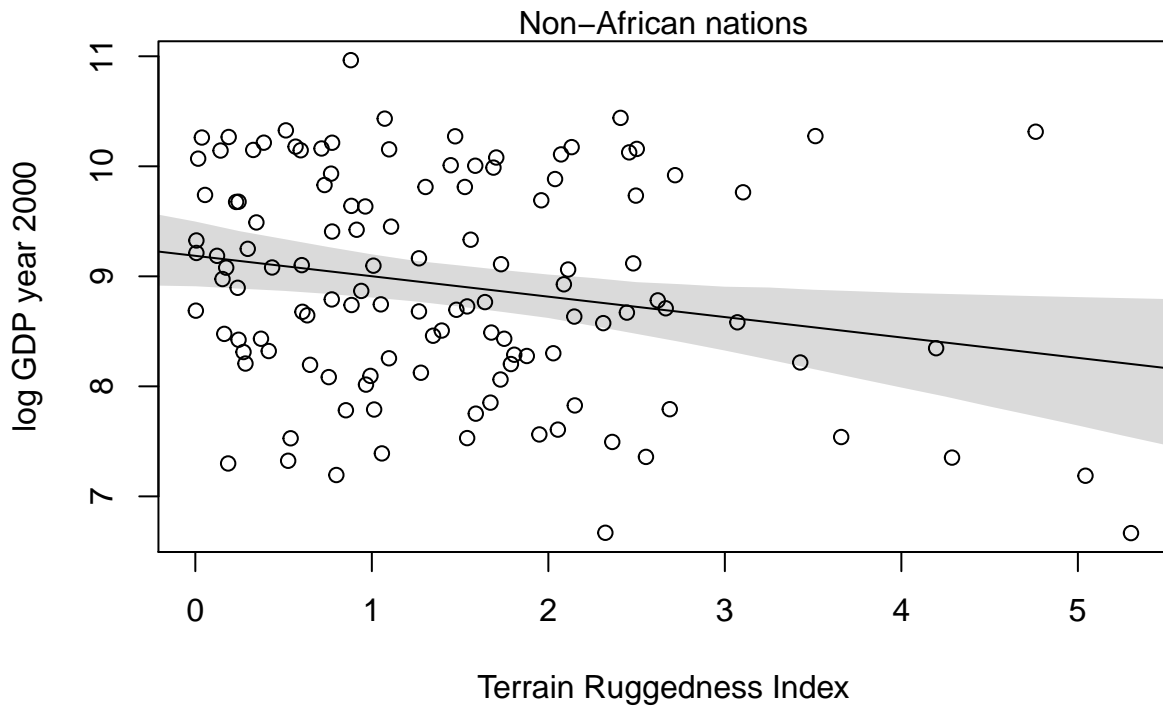
```
mtext( "African nations" , 3 )
```

```
lines( rugged.seq , mu.Africa.mean , col=rangi2 )
```

```
shade( mu.Africa.PI , rugged.seq , col=col.alpha(rangi2,0.3) )
```



```
# plot non-African nations with regression
d.AO <- dd[dd$cont_africa==0,]
plot( log(rgdppc_2000) ~ rugged , data=d.AO ,
      col="black" , ylab="log GDP year 2000" ,
      xlab="Terrain Ruggedness Index" )
mtext( "Non-African nations" , 3 )
lines( rugged.seq , mu.NotAfrica.mean )
shade( mu.NotAfrica.PI , rugged.seq )
```



## analysis

```
## R code 7.12
```

```
precis(m7.5)
```

```
##      Mean StdDev 5.5% 94.5%  
## a      9.18  0.14  8.97  9.40  
## bA     -1.85  0.22 -2.20 -1.50  
## bR     -0.18  0.08 -0.31 -0.06  
## bAR     0.35  0.13  0.14  0.55  
## sigma  0.93  0.05  0.85  1.01
```

```
## R code 7.13
```

```
post <- extract.samples( m7.5 )  
gamma.Africa <- post$bR + post$bAR*1  
gamma.notAfrica <- post$bR + post$bAR*0
```

```
## R code 7.14
```

```
mean( gamma.Africa)
```

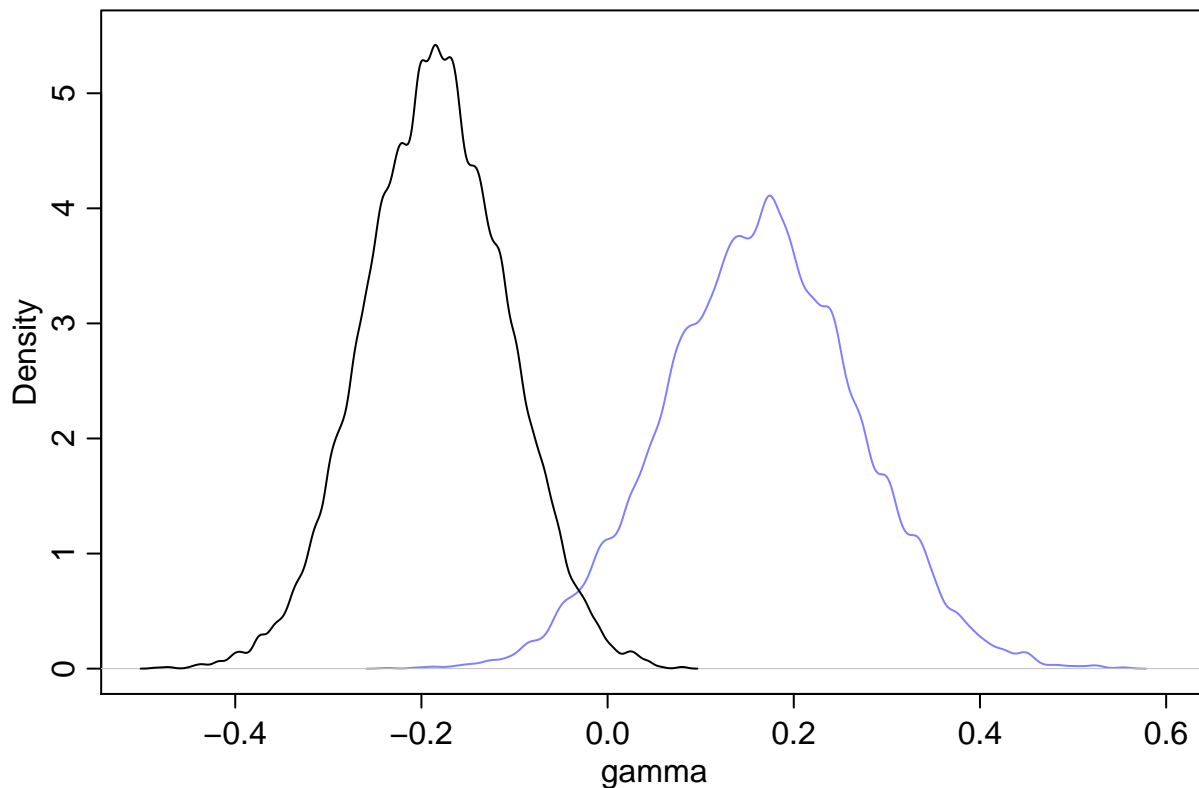
```
## [1] 0.1644571
```

```
mean( gamma.notAfrica )
```

```
## [1] -0.1847823
```

```
## R code 7.15
```

```
dens( gamma.Africa , xlim=c(-0.5,0.6) , ylim=c(0,5.5) ,  
      xlab="gamma" , col="rang12" )  
dens( gamma.notAfrica , add=TRUE )
```



```

## R code 7.16
diff <- gamma.Africa - gamma.notAfrica
sum( diff < 0 ) / length( diff )

## [1] 0.003

## R code 7.17
# get minimum and maximum rugged values
q.rugged <- range(dd$rugged)

# compute lines and confidence intervals
mu.ruggedlo <- link( m7.5 ,
  data=data.frame(rugged=q.rugged[1],cont_africa=0:1) )

## [ 100 / 1000 ]
[ 200 / 1000 ]
[ 300 / 1000 ]
[ 400 / 1000 ]
[ 500 / 1000 ]
[ 600 / 1000 ]
[ 700 / 1000 ]
[ 800 / 1000 ]
[ 900 / 1000 ]
[ 1000 / 1000 ]

mu.ruggedlo.mean <- apply( mu.ruggedlo , 2 , mean )
mu.ruggedlo.PI <- apply( mu.ruggedlo , 2 , PI )

mu.ruggedhi <- link( m7.5 ,
  data=data.frame(rugged=q.rugged[2],cont_africa=0:1) )

## [ 100 / 1000 ]
[ 200 / 1000 ]
[ 300 / 1000 ]
[ 400 / 1000 ]
[ 500 / 1000 ]
[ 600 / 1000 ]
[ 700 / 1000 ]
[ 800 / 1000 ]
[ 900 / 1000 ]
[ 1000 / 1000 ]

mu.ruggedhi.mean <- apply( mu.ruggedhi , 2 , mean )
mu.ruggedhi.PI <- apply( mu.ruggedhi , 2 , PI )

# plot it all, splitting points at median
med.r <- median(dd$rugged)
ox <- ifelse( dd$rugged > med.r , 0.05 , -0.05 )
plot( dd$cont_africa + ox , log(dd$rgdppc_2000) ,
  col=ifelse(dd$rugged>med.r,rangi2,"black") ,
  xlim=c(-0.25,1.25) , xaxt="n" , ylab="log GDP year 2000" ,
  xlab="Continent" )
axis( 1 , at=c(0,1) , labels=c("other","Africa") )
lines( 0:1 , mu.ruggedlo.mean , lty=2 )
shade( mu.ruggedlo.PI , 0:1 )
lines( 0:1 , mu.ruggedhi.mean , col=rangi2 )

```



```
shade( mu.ruggedhi.PI , 0:1 , col=col.alpha(rangi2,0.25) )
```

