

# Plotando Admixture

## Plotando Admixture

### Formatando los datos

Necesitamos de los archivos generados en Admixture: log, P y Q para cada valor de K. Con estos dos archivos vamos a plotar usando R con el paquete *tidyverse*.

Primero vamos a load los paquetes y setar el directorio

```
#library(tidyverse)
library(ggplot2)
#library(stringr)
setwd("~/Dropbox/Postdoc_Rosario/BiodiversityGenomics_Winter2021/dato_denovo/demographic/admixture/")
```

Ahora vamos a mirar los likelihoods de cada K

```
f <- list.files(path = ".", pattern = ".out")
lh_K <- data.frame(K = numeric(),
                  CV_error = numeric())

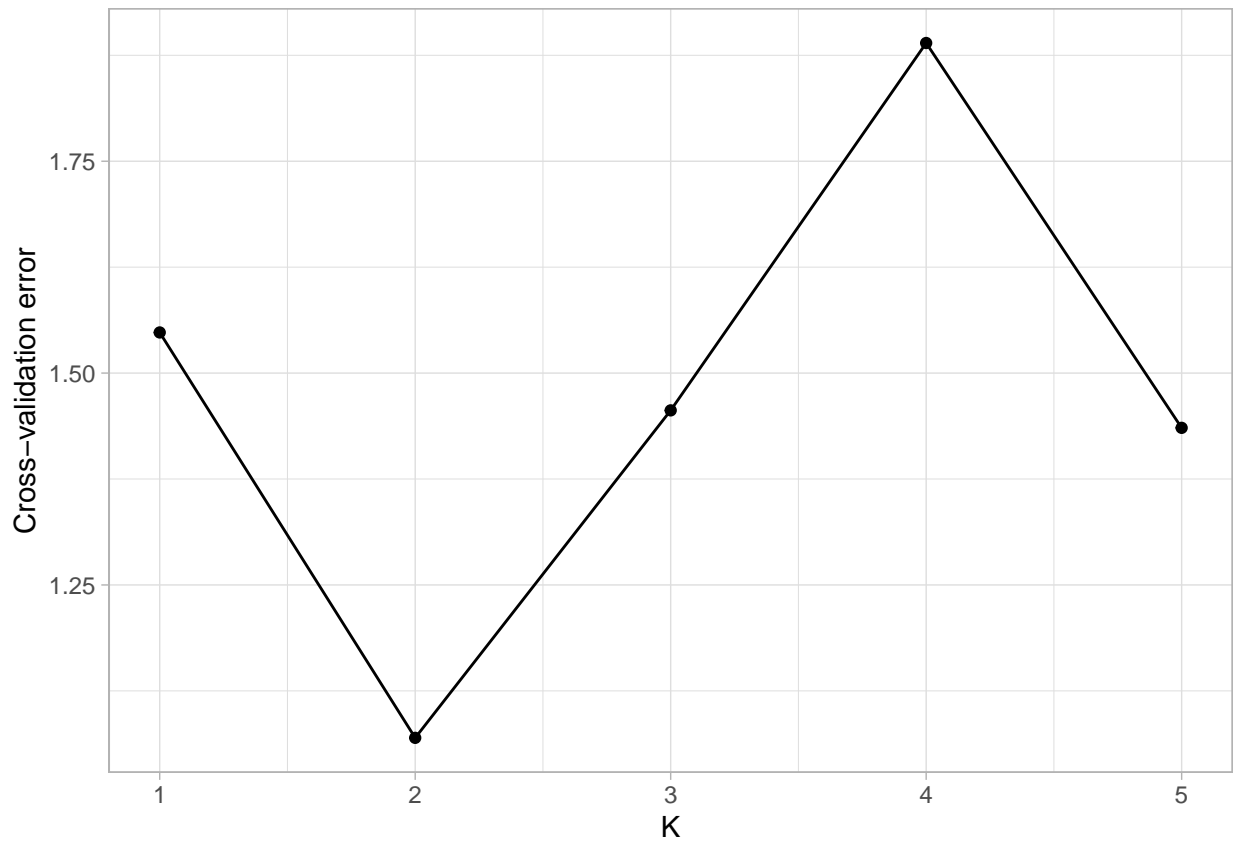
for(i in f){
  temp <- readLines(i)[grep("CV", readLines(i))]
  k <- as.integer(sub(".*?K=?(\\d+).*", "\\1", temp))
  cv <- as.numeric(sub(".*?): *(\\d+(?:\\.\\d+)).*", "\\1", temp))
  lh_K[nrow(lh_K) + 1,] <- c(k, cv)
}

print(lh_K)
```

```
##    K CV_error
## 1 1  1.54795
## 2 2  1.06957
## 3 3  1.45600
## 4 4  1.88951
## 5 5  1.43538
```

Plotando los likelihoods

```
a <- ggplot(lh_K, aes(x = K, y = CV_error)) +
  geom_point() +
  geom_line() +
  ylab("Cross-validation error") +
  xlab("K") +
  theme_light(); a
```



Ahora que ya descubrimos cual K es lo numero de clusters mas probable, vamos plotar la ancestralidad estimada por admixture para cada individuo. Pero antes de plotar necesitamos ler la matrix Q.

```
q <- read.table("populations_cor.2.Q")
barplot(t(as.matrix(q)), col = c("#ef8a62", "#91bfdb"), xlab = "Individual #", ylab = "Ancesrty")
```

