



۱ معرفی مدل رایت-فیشر بدون جهش

در مدل مذکور بدون در نظر گرفتن جهش، کسر تعداد هر ژن به سائز کل جمعیت به شکل زیر بیان می‌شود:

$$p_j = \frac{j}{2N}, \quad q_j = 1 - \frac{j}{2N} \quad (1)$$

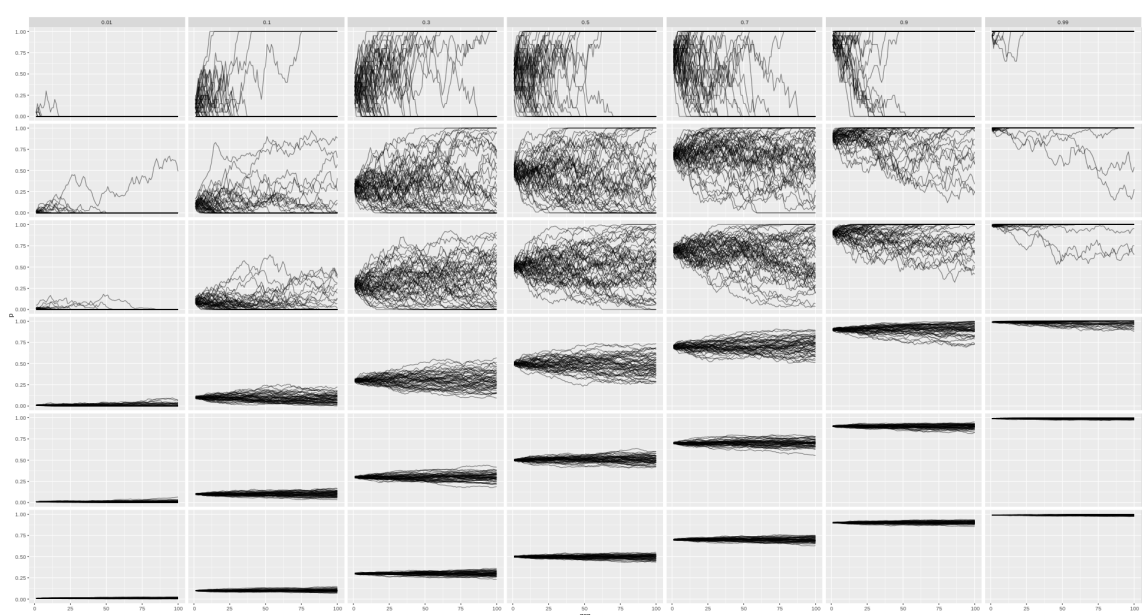
پیداست که ماتریس انتقال این فرآیند تصادفی نیز به صورت تابع چگالی احتمال دوجمله‌ای ساخته می‌شود:

$$P = (X_{n+1} = k | X_n = j) = p(i, j) = \binom{2N}{k} p_j^k q_j^{2N-k} \quad (2)$$

در این مدل دو حالت جاذب (Absorbing State) خواهیم داشت که شامل حالت‌هایی است که در آن‌ها یکی از ژن‌ها به طور کامل ژن غالب شده‌است (ژن دیگر منقرض شده‌است و در نسل‌های بعدی حضور نخواهد داشت).

۲ نتایج شبیه‌سازی بدون جهش

با ۵۰ بار شبیه‌سازی برای ۱۰۰ نسل برای جمعیت‌هایی با سائز متفاوت و مقادیر اولیه‌ی متفاوت کسر هر ژن از جمعیت دو مورد قابل توجه است. اول، همگرایی با افزایش سائز جمعیت مشخص‌تر می‌شود. دوم، حالات جاذب مدل با کسرهای اولیه‌ی نزدیک یک یا صفر مشخص‌تر اند و ادامه‌ی نسل‌شان با یک خط چسبیده به بالا یا پایین نمودار مشخص شده‌اند.



Initial values for p_j and population size are (.01, .1, .3, .5, .7, .9, .99) from left to right and (10, 50, 100, 1000, 5000, 15000) from top to bottom respectively.

```
library(ggplot2)
library(dplyr)
library(tidyr)
library(viridis)

wf_df <- data.frame()
sizes <- c(10, 50, 100, 1000, 5000, 15000)
starting_p <- c(.01, .1, .3, .5, .7, .9, .99)
n_gen <- 100
n_reps <- 50

for(N in sizes){
  for(p in starting_p){
    p0 <- p
    for(j in 1:n_gen){
      X <- rbinom(n_reps, 2*N, p)
      p <- X / (2*N)
      rows <- data.frame(replicate = 1:n_reps, N = rep(N, n_reps),
                        gen = rep(j, n_reps), p0 = rep(p0, n_reps),
                        p = p)
      wf_df <- bind_rows(wf_df, rows)
    }
  }
}

p <- ggplot(wf_df, aes(x = gen, y = p, group = replicate)) +
  geom_path(alpha = .5) + facet_grid(N ~ p0) + guides(colour=FALSE)
```

۳ معرفی مدل رایت-فیشر با جهش

حال برای نزدیک کردن مدل رایت-فیشر به واقعیت دو احتمال برای دو نوع جهش تعیین می‌کنیم. یک نوع جهش a به A و دیگری جهش A به a . با این کار طبیعتاً کسر تعداد هر ژن به کل جمعیت دستخوش تغییر مهمی می‌شود. اگر احتمال جهش نوع اول را با α_1 یک و احتمال جهش نوع دوم را با α_2 دو نشان دهیم، کسر تعداد ژن‌های A به کل جمعیت برابر زیر می‌شود:

$$p_j = \frac{j}{2N} (1 - \alpha_1) + (1 - \frac{j}{2N}) \alpha_2 \quad (3)$$

همچنین کسر تعداد ژن‌های a به کل جمعیت نیز به شکل زیر تغییر می‌کند:

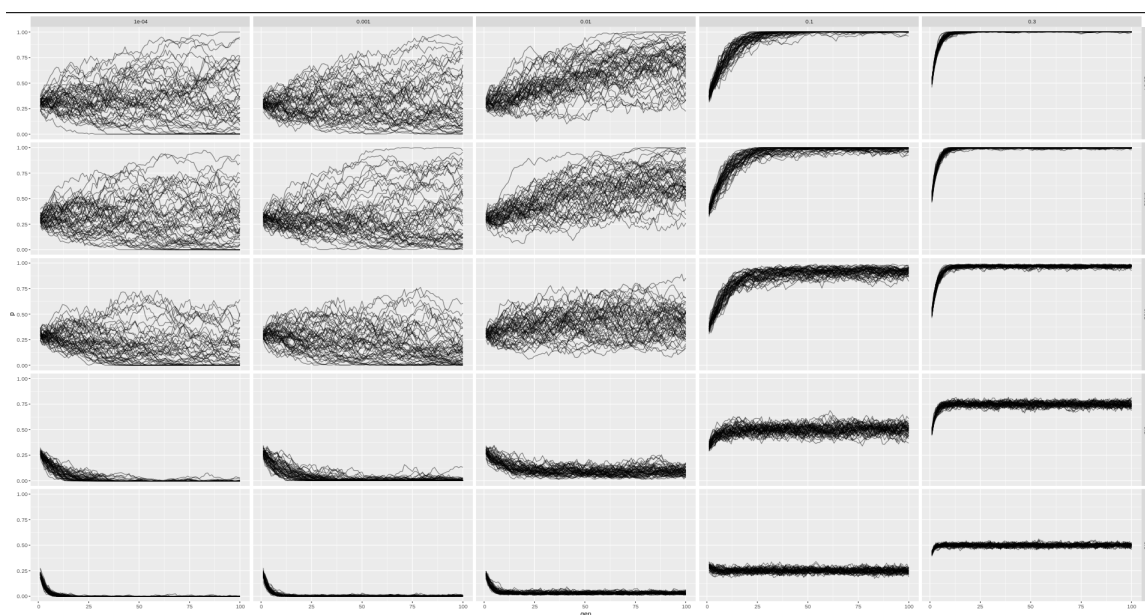
$$q_j = \frac{j}{2N} \alpha_1 + (1 - \frac{j}{2N}) (a - \alpha_2) \alpha_2 \quad (4)$$

با این حال تابع چگالی احتمال همچنان به صورت قبلی دوجمله‌ای باقی می‌ماند. (معادله‌ی ۲)

۴ نتایج شبیه‌سازی با جهش

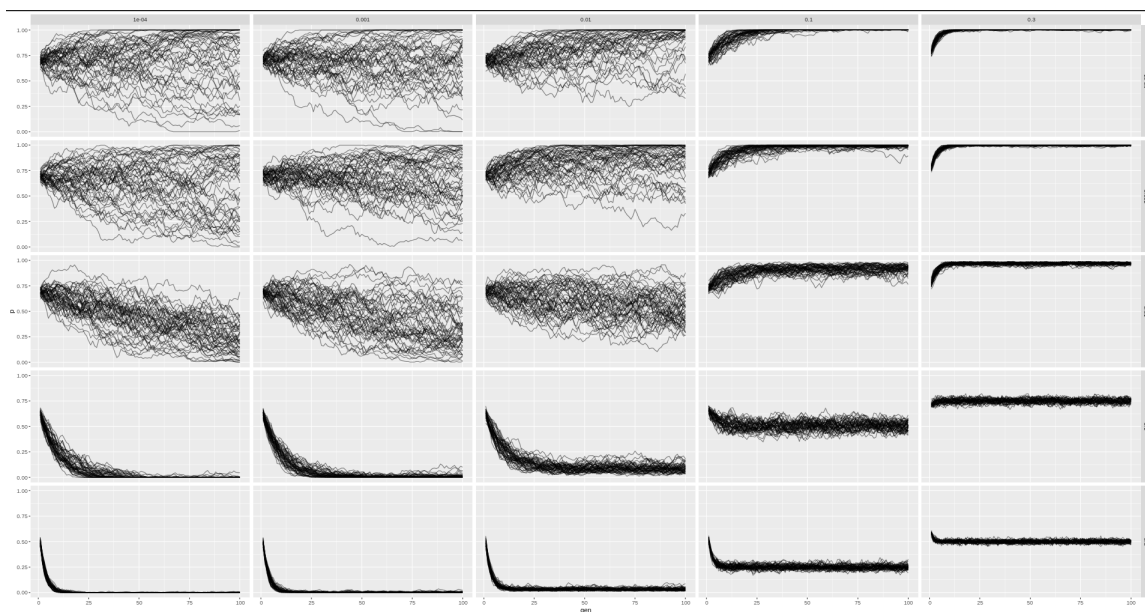
نتایج این بخش با توجه به اضافه شدن دو متغیر جدید در شبیه‌سازی مفصل‌تر است، برای همین در این بخش نمودارها با نظم جدیدی کنار هم قرار گرفته‌اند تا بتوان اثر متغیرهای احتمالات انواع جهش را مشاهده کرد. سائز جمعیت نمونه ۱۰۰ و ۱۰۰۰ در نظر گرفته شده است و کسر تعداد اولیه نیز فقط در حالات ۳۰ و ۷۰ درصد شبیه‌سازی شده‌اند.

جمعیت ۱۰۰ با کسر اولیه‌ی ۳۰ درصد:



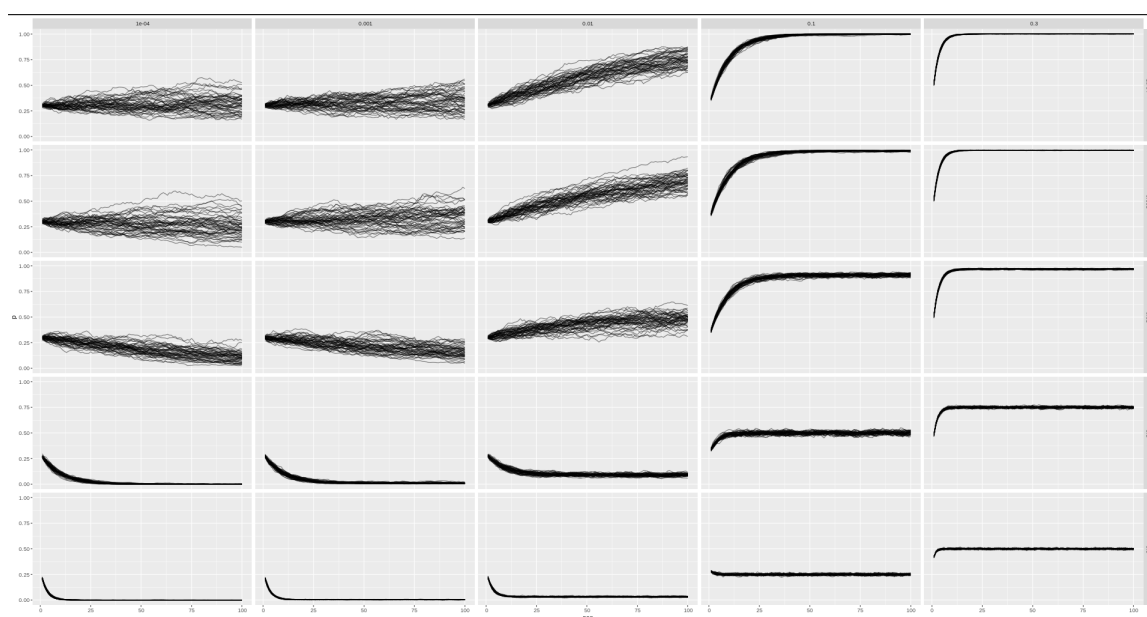
Initial values for both α_1 and α_2 are .0001, .001, .01, .1 and .3 from left to right and from top to bottom respectively.

جمعیت ۱۰۰ با کسر اولیه‌ی ۷۰ درصد:



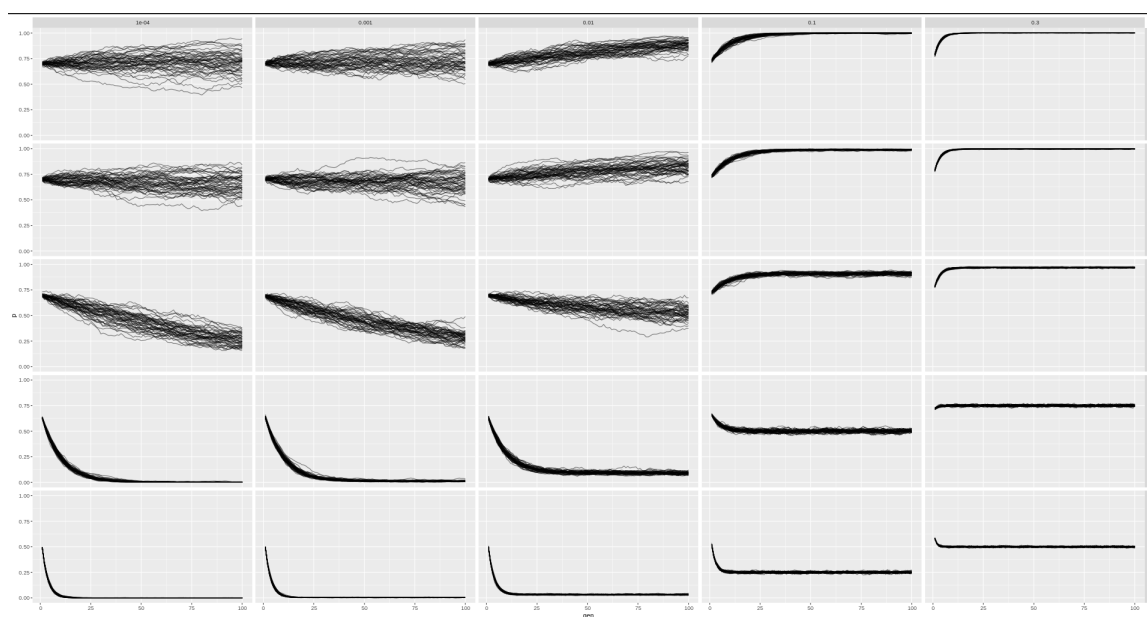
Initial values for both α_1 and α_2 are .0001, .001, .01, .1 and .3 from left to right and from top to bottom respectively.

جمعیت ۱۰۰۰ با کسر اولیه‌ی ۳۰ درصد:



Initial values for both α_1 and α_2 are .0001, .001, .01, .1 and .3 from left to right and from top to bottom respectively.

جمعیت ۱۰۰۰ با کسر اولیه‌ی ۷۰ درصد:



Initial values for both α_1 and α_2 are .0001, .001, .01, .1 and .3 from left to right and from top to bottom respectively.

کد شبیه‌سازی با جهش برای جمعیت ۱۰۰۰ و کسر اولیه‌ی ۷۰ درصد:

```
library(ggplot2)
library(dplyr)
library(tidyr)
library(viridis)

wf_df <- data.frame()
sizes <- 1000
starting_p <- .7
n_gen <- 100
n_reps <- 50
alpha1m <- c(.0001, .001, .01, .1, .3)
alpha2m <- c(.0001, .001, .01, .1, .3)

for(alpha1 in alpha1m){
  for(alpha2 in alpha2m){
    N <- sizes
    p <- starting_p
    p0 <- p
    for(j in 1:n_gen){
      X <- rbinom(n_reps, 2*N, p)
      p <- (X / (2*N)) * (1-alpha1) + (1 - (X / (2*N))) * alpha2
      rows <- data.frame(replicate = 1:n_reps, alpha1 = rep(alpha1, n_reps),
                        gen = rep(j, n_reps), alpha2 = rep(alpha2, n_reps),
                        p = p)
      wf_df <- bind_rows(wf_df, rows)
    }
  }
}

p <- ggplot(wf_df, aes(x = gen, y = p, group = replicate)) +
  geom_path(alpha = .5) + facet_grid(alpha1 ~ alpha2) + guides(colour=FALSE)
```

۵ ضمایم

دو قطعه کد برای دو شبیه‌سازی مدل رایت-فیشر در حالات انقراض و جهش ژن‌ها ضمیمه‌ی این گزارش شده‌اند.