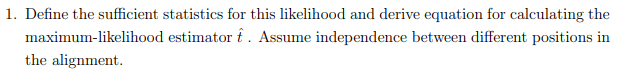
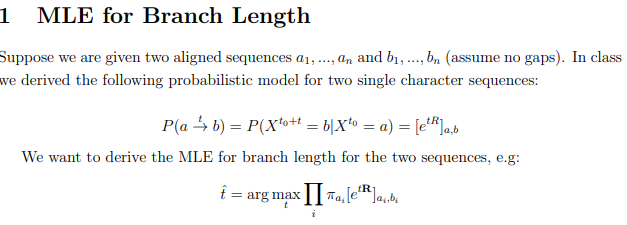
**אלגוריתמים בביולוגיה חישובית – תרגיל 4**



* 1. נרצה למצוא את ערך ה-t שממקסם את ה- log-likelihood, לשם כך נתחיל מלמצוא את הלוג של הביטוי, לאחר מכן נגזור ונשווה ל-0 למציאת המקסימום.

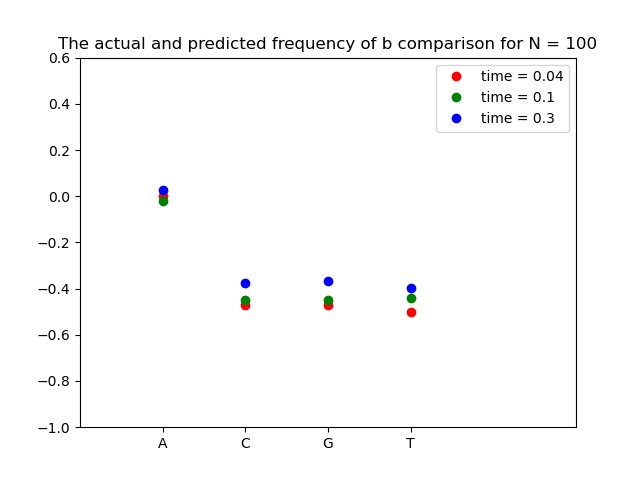
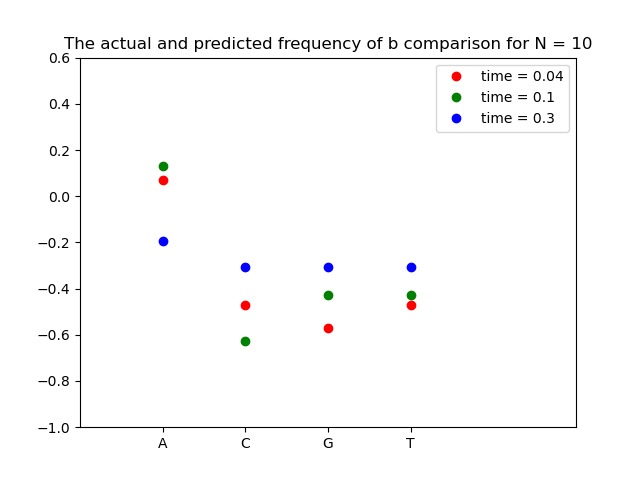
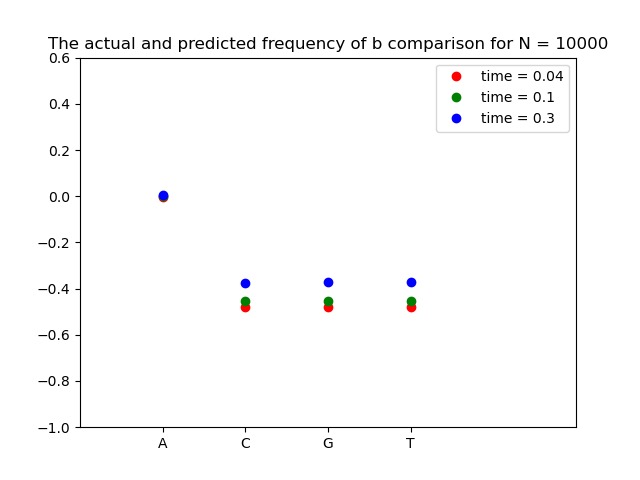
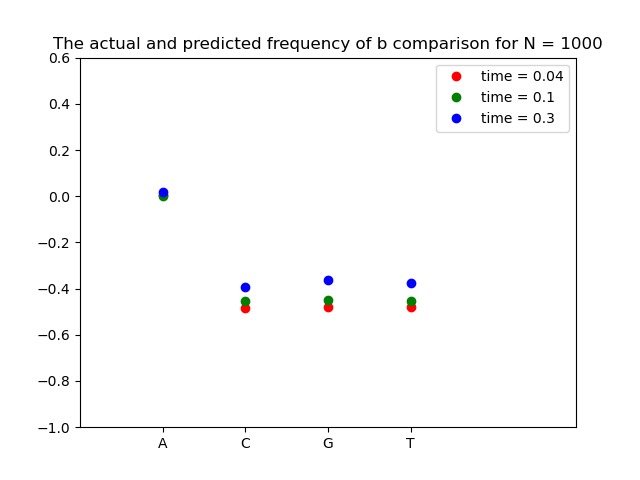
כעת נגדיר את הסטטיסטים בהתאם לחלקי המשוואה, כאשר נשים לב כי הביטוי האמצעי והימני קשורים זה לזה בכך שכמות האיברים שהם ביניהם היא n. נגדיר:

וכשנציב חזרה במשוואה:

נגזור:

ונשווה ל-0:



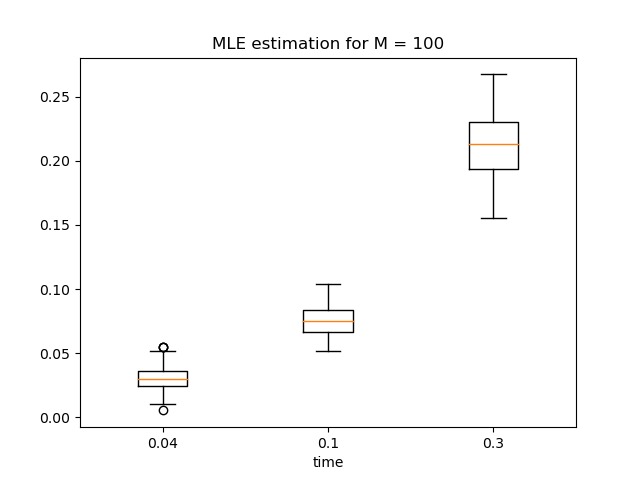


* 1. נביט בגרפים הבאים:

בגרף אנו יכולים לראות את אחוזי הדיוק שהמודל שלנו פלט כאשר בחרנו כמועמד את הנוקלאוטיד A ובדקנו את הסתברויות המעבר אל הנוקלאוטידים האחרים לפי הטבלה הסטציונרית שסופקה לנו.

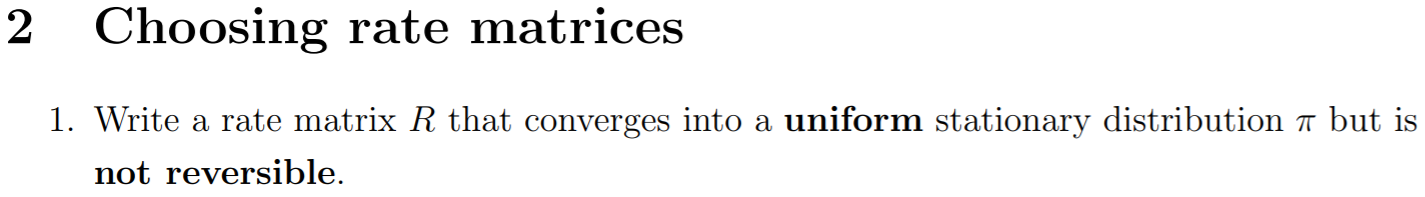
ניתן לראות שההבדלים בגרפים הם מספר החזרות, ובכל גרף דגמנו 3 פעמים בזמנים שונים.   
נתייחס תחילה לדגימות השונות בכל גרף. ניתן לראות כי ככל שנתנו למודל לרוץ יותר זמן כך התוצאות התכנסו לערך גבוהה יותר, כלומר המודל הצליח לחזות בהצלחה את הסתברות המעבר והנוקלאוטיד החזוי.  
כאשר נעבור מגרף לגרף נגדיל את מספר הדגימות עליהן המודל מתאמן, גם במקרה זה נראה שיפור של החיזויים והתכנסותם לערך יחיד.



* 1. 

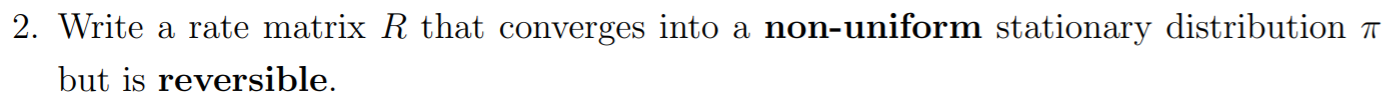
ביצענו 100 הרצות שבכל הרצה הגרלנו שני רצפים באורך  
500 כל אחד. עבור כל זמן t חישבנו את עומד ה-MLE,

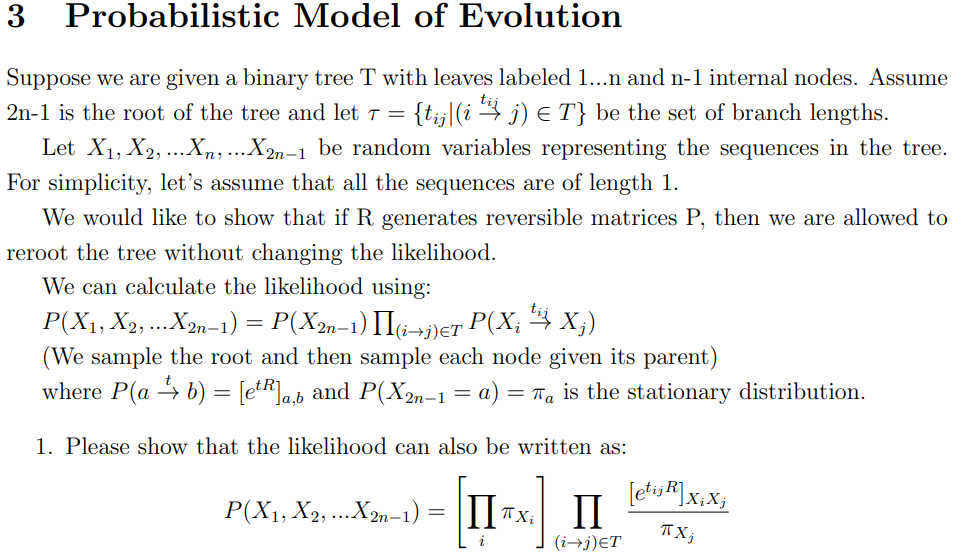
ניתן לראות שקיבלנו קירוב טוב לערכי ה-t שחזינו.  
בנוסף ניתן לראות שהשונות גדלה ככל שרצים יותר  
זמן על הרצף, תוצאה זו הגיונית מכיוון שככל שהזמן  
שנותנים לרצף להשתנות ולצבור "מוטציות" כך נראה  
יותר שוני בין הרצף ההתחלתי לסופי. דבר זה יכול להוות  
בעיה אם נרצה לעשות את הפעולה ההפוכה ולשחזר את  
העץ למציאת אב קדמון.



* 1. נרצה לבחור מטריצה שהיא לא רברסבילית, כלומר לא מקיימת את התכונה

אבל כן מתכנסת למטריצה אחידה, נבחר במטריצה:



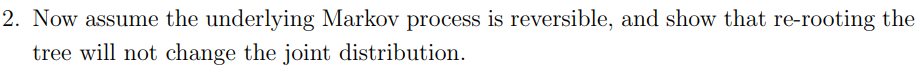


* 1. נפתח את המשוואה:

כאשר:

נכתוב את הביטויים לפי הגדרה.

נתון לנו שזהו עץ בינארי , לכן יש לו צלעות, כמספר הקודקודים פחות קודקוד השורש אליו אין אף צלע שנכנסת לכן לא נספור אותו.



* 1. נרצה לקחת את העץ "ולשתול" אותו מחדש עם שורש אחר, נעשה זאת ונשתמש בתוצאה מהסעיף הקודם לפתוח אותו בהתאם, נבחר את השורש החדש להיות :

כנדרש.

כאשר:

נוציא באופן מפורש את המעבר מהשורש החדש לשורש הקודם.

לצורך נוחות כתיבה נסמן:

השתמשנו בתכונת הרברסיביליות שלמדנו בשיעור: