ΣΤΟΧΑΣΤΙΚΕΣ ΑΝΕΛΙΞΕΙΣ
ΕΡΓΑΣΤΗΡΙΟ II

Πριν δοκιμάσετε τις ασκήσεις φροντίστε να έχετε εγκαταστημένο το πακέτο `python-tk`. Αν χρησιμοποιείτε κάποιο bundle ελέγξτε αν το `python-tk` είναι εγκαταστημένο, και αν όχι εγκαταστήστε το από τον διαχειριστή πακέτων. Αν χρησιμοποιείτε την εκδοχή του Linux δώστε την εντολή `sudo apt-get install python-tk` (ο κωδικός του administrator αν δεν τον έχετε άλλαξει είναι `markov`.)

Άσκηση 1 Σ' αυτήν την άσκηση θα δούμε πώς μπορούμε να κάνουμε γραφικές παραστάσεις με την Python. Κατεβάστε το πρόγραμμα `example_plot.py` και αποθηκεύστε το στον κατάλογο που θα δουλέψετε. Τρέξτε το πρόγραμμα. Το πρόγραμμα τυπώνει την γραφική παράσταση της συνάρτησης $x \mapsto 32x^3$ στο διάστημα (0,6) σε κανονική και λογαριθμική (με βάση το 2) κλίμακα.
α) Γιατί σε λογαριθμική κλίμακα βλέπουμε μια ευθεία;
β) Εκτιμήστε γραφικά την κλίση της ευθείας και το σημείο που τέμνει τον άξονα $y'y$.

Για τα επόμενα 2 ερωτήματα ίσως σας βοηθήσει να έχετε ανοίξει τον κώδικα με έναν επεξεργαστή κειμένου σε ένα τερματικό, και να εκτελείτε το πρόγραμμα από ένα άλλο τερματικό.

γ) Αλλάξτε την συνάρτηση σε $x \mapsto 8x^3$ και ξανατρέξτε το πρόγραμμα. Πώς αλλάζει το διάγραμμα σε λογαριθμική κλίμακα;
δ) Αλλάξτε την συνάρτηση σε $x \mapsto 8x^2$ και ξανατρέξτε το πρόγραμμα. Πώς αλλάζει το διάγραμμα σε λογαριθμική κλίμακα;

Κλείστε τώρα το παράθυρο γραφικών και αλλάξτε το πρόγραμμα ώστε να ξαναπάρετε την γραφική παράσταση της $x \mapsto 32x^3$. Αφαιρέστε το σύμβολο # που καθιστά σχολιασμό τις δύο γραμμές στο τέλος του προγράμματος

```
a,b = np.polyfit(newx,newy,1)

print "The fitted line is y=%.2f*x+%.2f " % (a,b)
```

Η πρώτη υπολογίζει τους συντελεστές της ευθείας $y = ax + b$ που προσαρμόζεται καλύτερα στα σημεία `(newx,newy)`. Η παράμετρος 1 δηλώνει ότι θέλουμε να προσαρμόσουμε ένα πολυώνυμο βαθμού 1, δηλαδή μια ευθεία. Η δεύτερη γραμμή απλά τυπώνει το αποτέλεσμα (με δύο δεκαδικά ψηφία) για τον χρήστη. Ξανατρέξτε το πρόγραμμα.

ε) Συμφωνεί το αποτέλεσμα για τα a, b με αυτό που βρήκατε στο ερώτημα (β);

Άσκηση 2 Κατεβάστε το πρόγραμμα `variance.py` και αποθηκεύστε το στον κατάλογο που θα δουλέψετε.

Το πρόγραμμα προσομοιώνει μια αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, 3, 4, 5\}$, και υπολογίζει με την μέθοδο Monte Carlo τον αναμενόμενο χρόνο επιστροφής στην κατάσταση 1, $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$, όπου

$$T_1^+ = \inf\{k > 0 : X_k = 1\}.$$

Η εκτίμηση για την $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$ λαμβάνεται προσομοιώνοντας την αλυσίδα N φορές, παίρνοντας N ανεξάρτητα δείγματα $t_1, \dots, t_N$ του χρόνου επιστροφής στο 1, και παίρνοντας τον μέσο όρο αυτών των δειγμάτων. Ο νόμος των μεγάλων αριθμών εγγυάται ότι ο μέσος όρος αυτών των N ανεξάρτητων δειγμάτων της τυχαίας μεταβλητής T_1^+ είναι για αρκετά μεγάλο N κοντά στην $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$ με μεγάλη πιθανότητα. Θα καλούμε την

$$E_N = \frac{t_1 + t_2 + \cdots + t_N}{N}$$

εκτιμήτρια Monte Carlo της αναμενόμενης τιμής $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$. Φυσικά η E_N είναι μια τυχαία μεταβλητή και αν ξανακάνουμε το πείραμα θα πάρουμε μια διαφορετική εκτίμηση.

α) Τρέξτε το πρόγραμμα μερικές φορές και δείτε πόσο διαφέρουν οι εκτιμήσεις που παίρνουμε κάθε φορά για την $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$. Επαναλάβετε για $N = 2^6, 2^7, \dots, 2^{12}$. Φαίνεται να διαφέρουν λιγότερο οι διαφορετικές εκτιμήσεις καθώς μεγαλώνει το N;

Σκοπός αυτής της άσκησης είναι να βρούμε υπολογιστικά πώς επηρεάζεται η διασπορά της εκτιμήτριας E_N από το πλήθος των επαναλήψεων N.

β) Φτιάξτε έναν βρόχο που θα κάνει $M = 30$ εκτιμήσεις της $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$ για κάθε δεδομένη τιμή του N και αποθηκεύστε αυτές τις M εκτιμήσεις στην λίστα mcestimates.

γ) Υπολογίστε την δειγματική μέση τιμή και την δειγματική διασπορά αυτών των εκτιμήσεων αφαιρώντας τον σχολιασμό από τις εντολές

```
sample_mean = float( sum(mcestimates) ) / M

squared_distance_from_mean = [ (e - sample_mean)**2 for e in mcestimates ]

sample_variance= float(sum ( squared_distance_from_mean )) / (M-1)
```

που βρίσκονται στο τέλος του προγράμματος.

δ) Φτιάξτε έναν βρόχο που κάνει τον παραπάνω υπολογισμό για $N = 2^6, 2^7, \ldots, 2^{14}$ και παραστήστε γραφικά πώς εξαρτάται η δειγματική μέση τιμή και η δειγματική τυπική απόκλιση από το N.

ε) Υπολογίστε θεωρητικά την $\mathbb{E}\big[T_1^+ \,|\, X_0 = 1\big]$. Συμφωνεί το αριθμητικό αποτέλεσμα που βρήκατε με την θεωρητική τιμή;

στ) Σχεδιάστε σε λογαριθμική κλίμακα πώς εξαρτάται η δειγματική τυπική απόκλιση της εκτιμήτριας Monte Carlo από το N. Ποια είναι η κλίση της ευθείας στο λογαριθμικό διάγραμμα; Συμφωνεί αυτό που βρήκατε με το κεντρικό οριακό θεώρημα;

Άσκηση 3 Σκοπός αυτής της άσκησης είναι να φτιάξουμε ένα animation μιας μαρκοβιανής αλυσίδας. Κατευάστε τον οδηγό `animationguide.pdf` και ακολουθήστε τα βήματα που περιγράφονται εκεί. Δείξτε 4 διαδοχικά στιγμιότυπα της αλυσίδας που χρησιμοποιήσατε.