Λύσεις των ασκήσεων

1. Αν $X \in \mathbb{N}$ είναι μια ΤΜ που παίρνει φυσικούς αριθμούς ως τιμές (και άρα έχει πάντα μέση τιμή, ενδεχομένως $+\infty$) δείξτε ότι

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \mathbb{P}[X \geq k].$$

Λύση: Έχουμε

$$\begin{aligned}\mathbb{E}[X] &= \sum_{n=1}^{\infty} n\mathbb{P}[X = n] \\ &= \sum_{n=1}^{\infty}\sum_{k=1}^{n} \mathbb{P}[X = n] \\ &= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} \mathbb{P}[X = n] \\ &= \sum_{k=1}^{\infty} \mathbb{P}[X \geq k].\end{aligned}$$

2. Σε μια πόλη έγινε ένας φόνος και η αστυνομία έχει καταλήξει ότι ο φονιάς είναι είτε ο Χ είτε ο Υ (και οι δύο είναι φυγάδες από το νόμο) και κατ' αρχήν η αστυνομία θεωρεί ότι είναι εξίσου πιθανοί ως δράστες. Αργότερα ήρθε η πληροφορία από το εργαστήριο ότι ο φονιάς έχει τύπο αίματος Α (10% του πληθυσμού έχει αυτόν τον τύπο αίματος).

Έρευνα στο αρχείο αποκαλύπτει αργότερα ότι ο Χ έχει τύπο αίματος Α (αλλά ουδέν είναι γνωστό για τον τύπο αίματος του Υ). Ποια η πιθανότητα ο Χ να είναι ο φονιάς μετά από τη νέα αυτή πληροφορία;

Λύση: Ας είναι H το ενδεχόμενο ο Χ να είναι ο φονιάς. Κατ' αρχήν έχουμε $\mathbb{P}[H] = \mathbb{P}[H^c] = 1/2$.

Αν E είναι το ενδεχόμενο ο Χ να έχει τύπο αίματος Α έχουμε

$$\mathbb{P}[E|H] = 1, \quad \mathbb{P}[E|H^c] = 1/10.$$

Με άλλα λόγια, αν ο Χ είναι ο φονιάς τότε σίγουρα έχει τύπο Α (αφού αυτό γνωρίζουμε για το φονιά), ενώ αν ο Χ δεν είναι ο φονιάς τότε η πιθανότητα να έχει τύπο Α είναι αυτή του γενικού πληθυσμού.

Όμως μαθαίνουμε ότι το E ισχύει, άρα πρέπει να υπολογίσουμε την πιθανότητα $\mathbb{P}[H|E]$. Χρησιμοποιώντας τον τύπο του Bayes και τις παραπάνω πληροφορίες παίρνουμε $\mathbb{P}[H|E] = \frac{10}{11}$.

3. Ένας τουρίστας με πολύ λίγη μνήμη θέλει να επισκεφθεί 4 πόλεις Α, Β, Γ, Δ. Το κάνει ως εξής: κάθε μέρα διαλέγει τυχαία μια από τις πόλεις (διαφορετική από την πόλη όπου βρίσκεται εκείνο το πρωί–θεωρείστε ότι το πρώτο πρωί βρίσκεται σε μια άλλη πόλη Ο) και την επισκέπτεται (δε θυμάται αν την έχει ήδη επισκεφθεί ή όχι). Κατά μέσο όρο πόσες μέρες θα του πάρει μέχρι να επισκεφθεί και τις 4 πόλεις;

Λύση: Αν είναι X ο αριθμός των ημερών τότε γράφουμε $X = X_1 + X_2 + X_3 + X_4$ όπου X_i είναι ο αριθμός των ημερών που περνάει από τότε που έχει επισκεφθεί συνολικά $i-1$ πόλεις έως τότε που έχει επισκεφθεί συνολικά i πόλεις.

Ισχύει προφανώς $X_1 = X_2 = 1$.

Ισχύει επίσης ότι η X_3 ακολουθεί γεωμετρική κατανομή με παράμετρο επιτυχίας $2/3$ και ότι η X_4 ακολουθεί επίσης γεωμετρική κατανομή αλλά με παράμετρο επιτυχίας $1/3$. Έπεται ότι $\mathbb{E}[X_3] = 3/2$ και $\mathbb{E}[X_4] = 3$, άρα $\mathbb{E}[X] = 1 + 1 + 3/2 + 3$.

4. Σε μια πόλη ενός εκατομμυρίου κατοίκων δύο άγνωστοι συναντιούνται. Ο καθένας έχει 500 γνωστούς στην πόλη (υποθέτουμε οι γνωστοί του καθενός είναι ένα τυχαίο σύνολο). Ποια η πιθανότητα ότι έχουν κοινό γνωστό;

Λύση: Μπορούμε χωρίς βλάβη της γενικότητας να υποθέσουμε ότι οι γνωστοί του πρώτου είναι οι υπ' αριθμόν 1 έως 500 κάτοικοι της πόλης. Εφόσον μας δίνεται ότι οι δύο που συναντιούνται μπορούμε να υποθέσουμε επίσης ότι ούτε ο ένας ούτε ο άλλος είναι ανάμεσα στους 1 έως 500, οπότε, χωρίς βλάβη της γενικότητας, μπορούμε να υποθέσουμε ότι τα δύο άτομα που συναντιούνται είναι οι υπ' αριθμόν 501 και 502 (με τους 1 έως 500) να είναι οι γνωστοί του 501).

Αν επιλέξουμε τους γνωστούς του 502 τυχαία ποια η πιθανότητα ότι κανείς από τους 500 γνωστούς του δεύτερου δεν είναι ανάμεσα στους 1 έως 500; Αυτή είναι

$$p = \frac{\binom{10^6-502}{500}}{\binom{10^6}{500}} = \frac{(10^6-502)\cdots(10^6-1001)}{10^6(10^6-1)\cdots(10^6-499)} = 0.7779245811309411,$$

άρα η πιθανότητα να υπάρχει κοινός γνωστός είναι η συμπληρωματική αυτής 0.2220754188690589.

Δεν υπάρχει ιδιαίτερα έξυπνος τρόπος υπολογισμού για το παραπάνω. Το απλούστερο είναι να γράφει κανείς ένα μικρό πρόγραμμα. Το παρακάτω πρόγραμμα σε Python 3 κάνει αυτή τη δουλειά:

```
p = 1
for i in range(500):
  p *= (1e6-502-i)/(1e6-i)
print(p)
```

5. Αν X, Y είναι ανεξάρτητες και ομοιόμορφες στο $[0,1]$ και $Z = XY$, βρείτε τη συνάρτηση κατανομής και τη συνάρτηση πυκνότητας της Z.

Λύση: Έχουμε

$$\begin{aligned} F_Z(z) &= \mathbb{P}[XY \le z] \\ &= \int_0^1 \int_0^{\min\{z/x,1\}} f_{(X,Y)}(x,y)\,dy\,dy \\ &= \int_0^1 \int_0^{\min\{z/x,1\}} 1\,dy\,dy \\ &= \int_0^1 \min\{z/x,1\}\,dx \\ &= z\int_z^1 \frac{dx}{x} + \int_0^z dx \\ &= z - z\log z, \end{aligned}$$

άρα, παραγωγίζοντας, έχουμε $f_Z(z) = -\log z \cdot \mathbb{1}(0 < z < 1)$.

6. Αν X, Y είναι ανεξάρτητες και ακολουθούν κανονική κατανομή με μέσους μ_X, μ_Y αντίστοιχα και διασπορές σ_X^2, σ_Y^2 δείξτε ότι και η $X+Y$ ακολουθεί κανονική κατανομή με μέσο $\mu_X + \mu_Y$ και διασπορά $\sigma_X^2 + \sigma_Y^2$.

Λύση: Δείτε την απόδειξη εδώ. Μπορείτε να υποθέσετε κατ' αρχήν ότι οι δύο μέσες τιμές είναι 0 και ότι οι δύο διασπορές είναι 1, ώστε να είναι απλούστερες οι πράξεις. Η όλη ουσία είναι να δείξει κανείς ότι άθροισμα ανεξαρτήτων κανονικών είναι και πάλι κανονική. Το ότι οι μέση τιμή του αθροίσματος είναι το άθροισμα των μέσων τιμών ισχύει πάντα και αφού οι μεταβλητές είναι ανεξάρτητες ισχύει επίσης ότι η διασπορά του αθροίσματος είναι το άθροισμα των διασπορών.

7. Μπορεί η X και η Y να είναι ομοιόμορφα κατανεμημένες στο $[0,1]$ και το ζεύγος (X, Y) να **μην** έχει πυκνότητα;

Λύση: Ναι, πάρτε X να είναι ομοιόμορφη στο $[0,1]$ και $Y = X$. Τότε υποσύνολα της ευθείας $L = \{(x,y) :\ x = y\}$ έχουν θετική πιθανότητα να πέσει μέσα η (X, Y). Αν υπήρχε πυκνότητα $f(x,y)$ του ζεύγους τότε θα έπρεπε λοιπόν να είναι 0 εκτός της ευθείας L (αλλιώς το ολοκλήρωμα της f σε κάποιο χωρίο D εκτός της ευθείας L θα έβγαινε θετικό, πράγμα αδύνατο). Όμως δε γίνεται μια συνάρτηση f να είναι 0 παντού εκτός της L και να έχει ολοκλήρωμα θετικό (πάρτε π.χ. προσέγγιση με Riemann αθροίσματα).

8. Το ζεύγος (X, Y) είναι ομοιόμορφα κατανεμημένο στο μοναδιαίο δίσκο $\left\{(x,y) :\ x^2 + y^2 \le 1\right\}$. Αν $R = \sqrt{X^2 + Y^2}$ είναι η απόσταση του τυχαίου σημείου από το $(0,0)$ βρείτε τη συνάρτηση κατανομής και τη συνάρτηση πυκνότητας της R.

Λύση: Βρίσκουμε πρώτα τη συνάρτηση κατανομής:

$$\begin{aligned} F_R(r) &= \mathbb{P}[R \le r] \\ &= \frac{\text{εμβαδό του δίσκου με ακτίνα } r}{\text{εμβαδό του μοναδιαίου δίσκου}} \\ &= \frac{\pi r^2}{\pi} \\ &= r^2, \end{aligned}$$

άρα $f_R(r) = 2r \cdot \mathbb{1}(0 \le r \le 1)$.

9. Το ζεύγος (X, Y) ανήκει πάντα στο μοναδιαίο δίσκο και έχει πυκνότητα η οποία είναι αυστηρά θετική μέσα στο μοναδιαίο δίσκο. Δείξτε ότι οι X και Y δε μπορεί να είναι ανεξάρτητες.

Λύση: Αν η $f_{(X,Y)}$ είναι αυστηρά θετική μέσα στο μοναδιαίο δίσκο τότε οι δύο περιθώριες είναι αυστηρά θετικές στο διάστημα $(-1, 1)$ (αφού είναι οι «προβολές» της $f_{(X,Y)}$ στους δύο άξονες). Αν οι X, Y είναι ανεξάρτητες τότε $f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y)$ και άρα η $f_{(X,Y)}(x, y)$ είναι αυστηρά θετική στο $[0, 1]^2$, πράγμα που δεν ισχύει γιατί μηδενίζεται εκτός του μοναδιαίου δίσκου.

10. Το ζεύγος (X, Y) ανήκει πάντα στο μοναδιαίο δίσκο και έχει πυκνότητα τέτοια ώστε
(1) Η ΤΜ $R = \sqrt{X^2 + Y^2}$ είναι ομοιόμορφα κατανεμημένη στο $[0, 1]$, και
(2) Η πυκνότητα του ζεύγους έχει κυκλική συμμετρία, δηλ. το $f_{(X,Y)}(x, y)$ εξαρτάται μόνο από την απόσταση του (x, y) από το $(0, 0)$.
Βρείτε την πυκνότητα $f_{(X,Y)}(x, y)$ και τις περιθώριες πυκνότητες.

Λύση: Το ότι η R έχει ομοιόμορφη κατανομή σημαίνει ότι $F_R(r) = r$ για $0 \le r \le 1$. Έχουμε, συμβολίζοντας με D_r το δίσκο με κέντρο το $(0, 0)$ και ακτίνα r:

$$\begin{aligned} r &= F_R(r) \\ &= \mathbb{P}\left[X^2 + Y^2 \le r^2\right] \\ &= \iint_{D_r} f_{(X,Y)}(x, y)\, dx\, dy \\ &= \int_0^{2\pi}\int_0^r f_{(X,Y)}(\rho)\rho\, d\rho\, d\theta \text{ (κάνοντας πολικές συν/μένες και χρησιμοποιώντας την συμμετρία)} \\ &= 2\pi \int_0^r f_{(X,Y)}(\rho)\rho\, d\rho. \end{aligned}$$

Παραγωγίζοντας ως προς r την ισότητα παίρνουμε

$$1 = 2\pi f_{(X,Y)}(r) r$$

άρα

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi r} = \frac{1}{2\pi\sqrt{x^2 + y^2}}.$$