Πιθανοτητες – Άνοιξη 2020-21

Λύσεις των ασκήσεων

1. Αποδείξτε τα παρακάτω για δύο οποιεσδήποτε ΤΜ X και Y:
(1) $\mathbb{V}ar\,[X \pm Y] = \mathbb{V}ar\,[X] + \mathbb{V}ar\,[Y] \pm 2\mathbb{C}ov\,(X, Y)$
(2) $\mathbb{C}ov\,(X + Y, Z) = \mathbb{C}ov\,(X, Z) + \mathbb{C}ov\,(Y, Z)$
(3) Αν $X_1, \ldots, X_n$ ανεξάρτητες τότε $\mathbb{C}ov\,(\sum_{j=1}^n a_j X_j, \sum_{j=1}^n b_j X_j) = \sum_{j=1}^n a_j b_j \mathbb{V}ar\,[X_j]$

Λύση: Αν T είναι μια ΤΜ γράφουμε για συντομία T' για την «κεντραρισμένη» T, δηλ. $T' = T - \mathbb{E}\,[T]$ (απαραίτητο είναι φυσικά να υπάρχει η $\mathbb{E}\,[T]$). Εύκολα βλέπουμε ότι η πράξη αυτή είναι γραμμική, δηλ. $(\lambda S + \mu T)' = \lambda S' + \mu T'$ για οποιεσδήποτε δύο σταθερές λ, μ και ΤΜ S, T που έχουν μέση τιμή.

Έτσι έχουμε $\mathbb{V}ar\,[T] = \mathbb{E}\left[T'^2\right]$ και $\mathbb{C}ov\,(S, T) = \mathbb{E}\,[S'T']$.

(1)

$$\mathbb{V}ar\,[X \pm Y] = \mathbb{E}\left[(X \pm Y)'^2\right] = \mathbb{E}\left[(X' \pm Y')^2\right] = \mathbb{E}\left[X'^2\right] + \mathbb{E}\left[Y'^2\right] \pm 2\mathbb{E}\,[X'Y'] = \mathbb{V}ar\,[X] + \mathbb{V}ar\,[Y] \pm 2\mathbb{C}ov\,(X, Y).$$

(2)

$$\mathbb{C}ov\,(X + Y, Z) = \mathbb{E}\,[X' + Y', Z'] = \mathbb{E}\,[X', Z'] + \mathbb{E}\,[Y', Z'] = \mathbb{C}ov\,(X, Z) + \mathbb{C}ov\,(Y, Z).$$

(3)

$$\mathbb{C}ov\,(\sum_{j=1}^n a_j X_j, \sum_{j=1}^n b_j X_j) = \mathbb{E}\left[\sum_{j=1}^n a_j X_j', \sum_{j=1}^n b_j X_j'\right] = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \mathbb{E}\left[X_i' X_j'\right] = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \mathbb{C}ov\,(X_i, X_j).$$

Αλλά αν $i \neq j$ οι X_i και X_j είναι ανεξάρτητες, άρα $\mathbb{C}ov\,(X_i, X_j) = 0$ οπότε από το παραπάνω άθροισμα απομένουν μόνο οι όροι με $i = j$, δηλ. μένει το άθροισμα

$$\sum_{j=1}^n a_j b_j \mathbb{C}ov\,(X_j, X_j) = \sum_{j=1}^n a_j b_j \mathbb{V}ar\,[X_i].$$

2. Η ΤΜ X παίρνει μόνο τις τιμές 1 και 2 και η ΤΜ Y παίρνει μόνο τις τιμές 0 και 1. Η κοινή τους πυκνότητα πιθανότητας είναι η εξής:

$$f_{(X,Y)}(1, 0) = 0.1,\ \ f_{(X,Y)}(1, 1) = 0.2,\ \ f_{(X,Y)}(2, 0) = 0.4,\ \ f_{(X,Y)}(2, 1) = 0.3.$$

Βρείτε πρώτα τις περιθώριες πυκνότητες f_X και f_Y, και μετά τις ποσότητες

$$\mathbb{E}\,[X]\,, \mathbb{E}\,[Y]\,, \sigma^2(X), \sigma^2(Y), \mathbb{C}ov\,(X, Y), \rho(X, Y).$$

($\rho(X, Y)$ είναι ο συντελεστής συσχέτισης των X, Y.)

Λύση: Έχουμε, κρατώντας την τιμή του X σταθερή και αθροίζοντας τις πιθανότητες για όλα τα δυνατά Y

$$f_X(1) = 0.3,\ \ f_X(2) = 0.7,$$

και, ομοίως εργαζόμενοι για την f_Y

$$f_Y(0) = f_Y(1) = 0.5.$$

Επίσης έχουμε

$$\begin{aligned} \mathbb{E}\,[X] &= 1 \times 0.3 + 2 \times 0.7 = 1.7, \\ \mathbb{E}\left[X^2\right] &= 1 \times 0.3 + 4 \times 0.7 = 3.1, \\ \sigma^2(X) &= \mathbb{E}\left[X^2\right] - \mathbb{E}\,[X]^2 = 3.1 - 2.89 = 0.21 \end{aligned}$$

και

$$\begin{aligned} \mathbb{E}\,[Y] &= 0 \times 0.5 + 1 \times 0.5 = 0.5, \\ \mathbb{E}\left[Y^2\right] &= 0 \times 0.5 + 1 \times 0.5 = 0.5, \\ \sigma^2(Y) &= \mathbb{E}\left[Y^2\right] - \mathbb{E}\,[Y]^2 = 0.5 - 0.25 = 0.25 \end{aligned}$$

Τέλος

$$\mathbb{C}ov\,(X, Y) = \mathbb{E}\,[XY] - \mathbb{E}\,[X]\,\mathbb{E}\,[Y] = 1 \cdot 0 \cdot 0.1 + 1 \cdot 1 \cdot 0.2 + 2 \cdot 0 \cdot 0.4 + 2 \cdot 1 \cdot 0.3 - 1.7 \cdot 0.5 = -0.05,$$

και

$$\rho(X, Y) = \frac{\mathbb{C}ov\,(X, Y)}{\sigma(X)\sigma(Y)} = \frac{-0.05}{0.4582575 \times 0.5} = -0.218217.$$

3. Σε ένα δοχείο έχουμε 3 κόκκινους και 2 μαύρους βώλους. Διαλέγουμε τυχαία δύο βώλους χωρίς επανάθεση. Έστω R ο αριθμός των κόκκινων βώλων που επιλέξαμε και B ο αριθμός των μαύρων. Υπολογίστε τις ποσότητες $\mathbb{E}\,[R]\,, \mathbb{E}\,[B]\,, \sigma^2(R), \sigma^2(B), \mathbb{C}ov\,(R, B)$.

Λύση: Έχουμε $R+B=2$ οπότε αρκεί να βρούμε την κατανομή του R και θα ξέρουμε τα πάντα (δηλ. την κοινή κατανομή των R, B).

Έχουμε $\mathbb{P}[R=0]=\frac{\binom{2}{2}}{\binom{5}{2}}=0.1$ (το κλάσμα προκύπτει επειδή επιλέγουμε και τους δύο βώλους από τους 2 μπλε), $\mathbb{P}[R=1]=\frac{3\cdot 2}{\binom{5}{2}}=0.6$ και $\mathbb{P}[R=2]=1-\mathbb{P}[R=0]-\mathbb{P}[R=1]=0.3$.

Άρα

$$\begin{aligned}\mathbb{E}[R]&=0\cdot 0.1+1\cdot 0.6+2\cdot 0.3=1.2,\\ \mathbb{E}\left[R^2\right]&=0\cdot 0.1+1\cdot 0.6+4\cdot 0.3=1.8,\\ \sigma^2(R)&=\mathbb{E}\left[R^2\right]-\mathbb{E}[R]^2=1.8-1.2^2=0.36,\end{aligned}$$

απ' όπου προκύπτει

$$\begin{aligned}\mathbb{E}[B]&=2-\mathbb{E}[R]=0.8,\\ \mathbb{E}\left[B^2\right]&=\mathbb{E}\left[(2-R)^2\right]=\mathbb{E}\left[R^2\right]+4-4\mathbb{E}[R]=1.8+4-4\cdot 1.2=1\\ \sigma^2(B)&=\mathbb{E}\left[B^2\right]-\mathbb{E}[B]^2=1-0.8^2=0.36,\end{aligned}$$

με την $\sigma^2(B)$ να μπορεί να προκύψει από την $\sigma^2(R)$ και χωρίς καμία πράξη αφού

$$\sigma^2(B)=\sigma^2(2-R)=\sigma^2(R-2)=\sigma^2(R).$$

Τέλος έχουμε

$$\mathbb{C}ov(R,B)=\mathbb{C}ov(R,2-R)=\mathbb{C}ov(R,2)-\mathbb{C}ov(R,R)=0-\sigma^2(R)=-0.36.$$

4. Έστω π μια τυχαία μετάθεση του συνόλου $S=\{1,2,\ldots,n\}$ (με πιο απλά λόγια, είναι ένας τυχαίος τρόπος να γράψουμε τα στοιχεία αυτά στη σειρά, κι όλοι οι τρόποι είναι εξίσου πιθανοί). Ας είναι X ο αριθμός των σταθερών σημείων της π, δηλ. $X=|\{x\in S:\pi(x)=x\}|$. Βρείτε τη μέση τιμή $\mathbb{E}[X]$ και τη διασπορά $\sigma^2(X)$.

Λύση: Γράφουμε $X=\sum_{j=1}^n \mathbb{1}(\pi(j)=j)$ απ' όπου προκύπτει

$$\mathbb{E}[X]=\sum_{j=1}^n \mathbb{P}[\pi(j)=j],$$

και, λόγω συμμετρίας,

$$\mathbb{E}[X]=n\mathbb{P}[\pi(1)=1].$$

Όμως όλες οι εικόνες $\pi(1)$ είναι εξίσου πιθανές (και πάλι λόγω συμμετρίας) και άρα $\mathbb{P}[\pi(1)=1]=1/n$ οπότε $\mathbb{E}[X]=1$.

Έχουμε επίσης

$$X^2=\sum_{i,j=1}^n \mathbb{1}(\pi(i)=i,\ \pi(j)=j)=\sum_{i=1}^n \mathbb{1}(\pi(i)=i)+\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbb{1}(\pi(i)=i,\ \pi(j)=j),$$

άρα (εκμεταλλευόμενοι και πάλι τη συμμετρία σε διάφορα σημεία)

$$\mathbb{E}\left[X^2\right]=\sum_{i=1}^n \mathbb{P}[\pi(1)=1]+\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbb{P}[\pi(1)=1,\ \pi(2)=2]=n\frac{1}{n}+n(n-1)\frac{1}{n(n-1)}=1+1=2,$$

όπου έχουμε $\mathbb{P}[\pi(1)=1,\ \pi(2)=2]=\frac{(n-2)!}{n!}=\frac{1}{n(n-1)}$ γιατί υπάρχουν ακριβώς $(n-2)!$ μεταθέσεις (αυτές των συμβόλων $3,4,\ldots,n$) που ικανοποιούν τη συνθήκη $\pi(1)=1,\ \pi(2)=2$.

5. Ας είναι A και B δύο τυχαία υποσύνολα του $\{1,2,\ldots,n\}$ μεγέθους k. Βρείτε τη μέση τιμή $\mathbb{E}[|A\cap B|]$.

Λύση: Έχουμε

$$\mathbb{E}[|A\cap B|]=\sum_{i=1}^n \mathbb{P}[i\in A\cap B]=\sum_{i=1}^n \mathbb{P}[i\in A, i\in B]=\sum_{i=1}^n \mathbb{P}[i\in A]\mathbb{P}[i\in B]=n\mathbb{P}[1\in A]^2,$$

και, αφού ισχύει

$$\mathbb{P}[1\in A]=\frac{\binom{n-1}{k-1}}{\binom{n}{k}}=\frac{k}{n},$$

παίρνουμε

$$\mathbb{E}[|A\cap B|]=n\frac{k^2}{n^2}=\frac{k^2}{n}.$$

6. Έχουμε n άτομα των οποίων οι ημέρες γενεθλίων είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες στις 365 μέρες του έτους. Ας είναι X το πλήθος των ζευγών ατόμων που έχουν τα ίδια γενέθλια. Βρείτε τις μέσες τιμές $\mathbb{E}[X], \mathbb{E}[X^2]$.

Λύση: Αν γράψουμε $b(i) \in \{1, 2, \ldots, 365\}$ για την ημέρα των γενεθλίων του ατόμου i έχουμε, για $i \neq j$ και λόγω της ανεξαρτησίας $\mathbb{P}[b(i) = b(j)] = \frac{1}{365}$.

Έτσι

$$\mathbb{E}[X] = \sum_{\substack{i,j=1 \\ i<j}}^{n} \mathbb{P}[b(i) = b(j)] = \frac{1}{365}\binom{n}{2} = \frac{n(n-1)}{2 \cdot 365}.$$

Επίσης

$$\mathbb{E}[X^2] = \mathbb{E}\left[\left(\sum_{\substack{i,j=1 \\ i<j}}^{n} \mathbb{1}(b(i) = b(j))\right)^2\right] = \sum_{\substack{i,j=1 \\ i<j}}^{n} \sum_{\substack{k,l=1 \\ k<l}}^{n} \mathbb{P}[b(i) = b(j),\ b(k) = b(l)]$$

Για να υπολογίσουμε το τελευταίο άθροισμα θα πρέπει να διαχωρίσουμε τους προσθετέους σε διάφορες κατηγορίες. Αυτές είναι οι εξής:

Περίπτωση 1: Όλα τα i, j, k, l είναι διαφορετικά. Σε αυτή την περίπτωση έχουμε $\mathbb{P}[b(i) = b(j),\ b(k) = b(l)] = \frac{1}{365^2}$ λόγω ανεξαρτησίας και το πλήθος των όρων που είναι αυτή την κατηγορία είναι $\binom{n}{2}\binom{n-2}{2} = \frac{1}{4}n(n-1)(n-2)(n-3)$. Άρα η συνεισφορά της περίπτωσης αυτής στο παραπάνω άθροισμα είναι

$$S_1 = \frac{1}{4 \cdot 365^2} n(n-1)(n-2)(n-3).$$

Περίπτωση 2: $i = k, j = l$ (και πάντα φυσικά ισχύει $i < j, k < l$). Σε αυτή την περίπτωση έχουμε

$$\mathbb{P}[b(i) = b(j),\ b(k) = b(l)] = \frac{1}{365},$$

και το πλήθος των όρων αυτής της κατηγορίας είναι $\binom{n}{2}$. Η συνεισφορά της περίπτωσης αυτής στο παραπάνω άθροισμα είναι

$$S_2 = \frac{1}{2 \cdot 365} n(n-1).$$

Περίπτωση 3: υπάρχει ακριβώς ένα κοινό άτομο στα δύο ζεύγη $\{i, j\}$ και $\{k, l\}$. Με άλλα λόγια $|\{i, j\} \cap \{k, l\}| = 1$. Σε αυτή την περίπτωση η συνθήκη $b(i) = b(j),\ b(k) = b(l)$ είναι σα να ζητάμε τα τρία άτομα που απαρτίζουν το σύνολο $\{i, j\} \cup \{k, l\}$ να έχουν τα ίδια γενέθλια. Λόγω ανεξαρτησίας η πιθανότητα αυτή είναι $\frac{1}{365^3}$ και ο αριθμός των όρων του αθροίσματος που συμπεριλαμβάνεται σε αυτή την περίπτωση δεν είναι παρά το συμπλήρωμα των δύο προηγούμενων περιπτώσεων, δηλ. είναι

$$\binom{n}{2}^2 - \frac{1}{4}n(n-1)(n-2)(n-3) - \binom{n}{2} = n(n-1)(n-2),$$

και άρα η συνεισφορά αυτής της περίπτωσης είναι

$$S_3 = \frac{n(n-1)(n-2)}{365^3}.$$

Παίρνουμε τέλος

$$\mathbb{E}[X^2] = S_1 + S_2 + S_3.$$