ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ, ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΙΘΑΝΟΤΗΤΕΣ, ΣΤΑΥΡΟΣ ΤΟΥΜΠΗΣ
ΤΕΛΙΚΗ ΕΞΕΤΑΣΗ, ΙΟΥΝΙΟΣ 2020
ΟΝΟΜΑ ΦΟΙΤΗΤΗ: ....

Οδηγίες

1. **Συμπληρώστε το όνομά σας άνω, και παραδώστε το παρόν με τις λύσεις**.
2. Διάρκεια εξέτασης: 100 ΛΕΠΤΑ
3. Απαγορεύεται η αναχώρηση από την αίθουσα πριν την συμπλήρωση 30λέπτου.
4. Απαγορεύεται η χρήση υπολογιστή χειρός. Απαγορεύεται η χρήση κινητού, και ως υπολογιστή χειρός.
5. Οι λύσεις πρέπει να γραφούν αποκλειστικά στην παρεχόμενη κόλλα.
6. Μπορείτε να χρησιμοποιείτε μολύβι ή/και στυλό οποιουδήποτε χρώματος εκτός από κόκκινο.
7. Οι λύσεις πρέπει να είναι το κατά δυνατόν αναλυτικές. **Πρέπει να φαίνονται όλα τα ενδιάμεσα βήματα στους υπολογισμούς. Τοποθετήστε τα τελικά αποτελέσματα εντός πλαισίου.**
8. Ξεκινήστε από αυτές τις ασκήσεις που ξέρετε και/ή δίνουν πολλές μονάδες.
9. Προχωρήστε κάθε άσκηση όσο μπορείτε! Θα δοθούν μονάδες για ασκήσεις λυμένες εν μέρει.
10. Στις τελικές απαντήσεις αρκεί να δοθούν μαθηματικές εκφράσεις, δεν χρειάζεται να κάνετε πράξεις. Μπορείτε να χρησιμοποιήσετε αιτιολογημένες προσεγγίσεις.
11. Σύμφωνα με τον κανονισμό εξετάσεων, σε περίπτωση που διαπιστωθεί (κατά τη διάρκεια της εξέτασης είτε της διόρθωσης) αντιγραφή είτε απόπειρα αντιγραφής, θα ενημερωθούν τα αρμόδια όργανα του ιδρύματος.

ΘΕΜΑΤΑ

1. **(Αεροπλάνο)** Σε ένα αεροπλάνο επιβαίνουν 40 Βρετανοί, 60 Γάλλοι, και 80 Ρουμάνοι. Καθένας από τους Βρετανούς είναι φορέας του κορωνοϊού με πιθανότητα 20%, ενώ οι αντίστοιχες πιθανότητες για τους Γάλλους και τους Ρουμάνους είναι 10% και 5%.
   (α΄) *(1 μονάδα)* Επιλέγεται ένα άτομο στην τύχη, χωρίς κάποια προτίμηση, και προκύπτει πως είναι φορέας. Ποια είναι η πιθανότητα να είναι Ρουμάνος;
   (β΄) *(1.5 μονάδα)* Επιλέγονται 4 άτομα στην τύχη, χωρίς προτίμηση στο συνδυασμό. Ποια η πιθανότητα να είναι όλοι Ρουμάνοι και επιπλέον ακριβώς 3 από αυτούς να είναι φορείς του κορωνοϊού;
2. **(Υπέρβαση χωρητικότητας)** *(1 μονάδα)* Σε μια εξέταση πιθανοτήτων έχουν λάβει οδηγία να καθίσουν σε μια αίθουσα 22 άτομα. Καθένα από αυτά τα άτομα θα έρθει με πιθανότητα $\frac{3}{4}$. Η αίθουσα χωρά να καθίσουν μόλις 20. Ποια είναι η πιθανότητα p τελικά να υπάρξει τουλάχιστον ένα άτομο που δεν θα βρει θέση να καθίσει σε εκείνη την αίθουσα;
3. **(Μετάδοση κορωνοϊού)** *(1 μονάδα)* Σε μια κοινωνική εκδήλωση συμμετέχουν 20 άτομα, εκ των οποίων 5 είναι φορείς του κορωνοϊού. Κάθε ένας από τους φορείς θα μεταδώσει, μέχρι το τέλος της εκδήλωσης, τον κορωνοϊό σε κάθε έναν από τους υπόλοιπους 15 με πιθανότητα 0.1. Ποια είναι η μέση τιμή των ατόμων που θα έχουν γίνει φορείς μέχρι το τέλος της εκδήλωσης; Εννοείται πως αρκεί μια μόνο μετάδοση για να γίνει κάποιος φορέας, και πως οι μεταδόσεις είναι ανεξάρτητες.

4. **(Τοστιέρα-αποστειρωτής)** Ένας διδάσκων πιθανοτήτων αποστειρώνει γραπτά της τελικής εξέτασης στην τοστιέρα του σπιτιού του. Προκειμένου να αποστειρωθεί πλήρως ένα γραπτό, πρέπει να παραμείνει στην τοστιέρα για ένα χρόνο X ο οποίος είναι Τ.Μ. με πυκνότητα

$$f(x) = \begin{cases} C(x-1)e^{-x}, & x \geq 1, \\ 0, & x < 1, \end{cases}$$

όπου το C είναι μια άγνωστη σταθερά.

(α΄) *(1 μονάδα)* Προσδιορίστε την τιμή της σταθεράς C.

(β΄) *(1 μονάδα)* Να δώσετε ένα τύπο για την συνάρτηση κατανομής πιθανότητας $F_X(x)$ της X.

(γ΄) *(1 μονάδα)* Καθώς ο διδάσκων δεν έχει στη διάθεσή του άπειρο χρόνο, κρατά στην τοστιέρα κάθε γραπτό για χρόνο θ ο οποίος εξασφαλίζει ότι το γραπτό θα έχει αποστειρωθεί πλήρως με πιθανότητα 99%. Βρείτε μια εξίσωση που πρέπει να ικανοποιεί το θ. (Μην επιχειρήσετε να την λύσετε ως προς θ, καθώς δεν μπορεί να γραφτεί σε κλειστή μορφή χρησιμοποιώντας απλές συναρτήσεις.)

(δ΄) *(1 μονάδα)* Για τον χρόνο θ του προηγούμενου σκέλους, ποια ακριβώς είναι η κατανομή των γραπτών που ΔΕΝ έχουν αποστειρωθεί πλήρως, αν υποθέσουμε ότι υπάρχουν 400 γραπτά και οι χρόνοι αποστείρωσης είναι ανεξάρτητοι μεταξύ τους; Ποια είναι επακριβώς η πιθανότητα να μην έχουν αποστειρωθεί ακριβώς 10 γραπτά; Μπορείτε να γράψετε μια προσέγγιση για αυτή την πιθανότητα;

5. **(Δηλώσεις συμμετοχής)** *(1.5 μονάδα)* Σε ένα μάθημα πιθανοτήτων έχουν δηλώσει συμμετοχή 600 άτομα. Καθένας όμως από αυτούς θα προσέλθει στο μάθημα με πιθανότητα $\frac{2}{3}$, ανεξάρτητα από τους υπόλοιπους. Να υπολογίσετε προσεγγιστικά την πιθανότητα να έρθουν, τελικά, πάνω από 420 άτομα.

ΤΥΠΟΛΟΓΙΟ ΠΙΘΑΝΟΤΗΤΩΝ

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad (A \cup B)' = A' \cap B', \quad (A \cap B)' = A' \cup B',$$
$$P(A) \geq 0 \, \forall A \in \mathcal{F}, \quad P(\Omega) = 1, \quad A_i \cap A_j = \emptyset \Rightarrow P({}_1 \cup {}_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots,$$
$$P(A') = 1 - P(A), \quad P(\emptyset) = 0, \quad P(A) \leq 1, \quad A \subseteq B \Rightarrow P(A) \leq P(B), \quad P(A \cap B') = P(A) - P(A \cap B),$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \; P(A \cup B) \leq P(A) + P(B), \; P(A \cup B \cup C) = P(A) + P(A' \cap B) + P(A' \cap B' \cap C),$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C), \; P(A) = P(A \cap B) + P(A \cap B').$$

Διατάξεις k αντικ. από N: $\frac{N!}{(N-k)!}$, Συνδυασμοί k αντικ. από N: $\binom{N}{k} = \frac{N!}{k!(N-k)!}$,
Επαναληπτικές διατάξεις μήκους k από αντικ.: N^k, Επαναληπτικοί συνδυασμοί k αντικ. από N: $\binom{+k-1}{k} = \binom{+k-1}{-1}$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(A_1 A_2 \ldots A_n) = P(A_1) P(A_2|A_1) P(A_3|A_2 A_1) \ldots P(A_n|A_1 A_2 \ldots A_{n-1}),$$
$$P() = P(A|B)P(B) + P(A|B')P(B'), \quad P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B')P(B')},$$
$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N) \quad (\{B_i\} \text{ διαμέριση}),$$
$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N)} \quad (\{B_i\} \text{ διαμέριση}).$$
$$A, B \text{ ανεξάρτητα} \Leftrightarrow P(AB) = P(A)P(B) \Leftrightarrow P(B) = P(B|A) \Leftrightarrow P(A) = P(A|B).$$

$$p_X(x) = P(X = x) = P(\{\omega : X(\omega) = x\}) \, \forall x \in S_X, \quad F_X(x) = P(X \leq x), \quad \lim_{t \to x^-} F(t) = P(X < x), \quad E(X) = \sum_{x \in S_X} x p_X(x),$$
$$E(g(X)) = \sum_{x \in S_X} g(x) p_X(x), \quad E(aX + b) = aE(X) + b, \quad E(\sum_{i=1}^n a_i g_i(X)) = \sum_{i=1}^n a_i E(g_i(X)),$$
$$\sigma^2 = \mathrm{VAR}(X) = E\big[(X - \mu)^2\big] = E(X^2) - (E(X))^2, \quad \mathrm{VAR}(aX + b) = a^2 \mathrm{VAR}(X).$$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad \sum_{k=0}^{\infty} k x^k = \frac{x}{(1-x)^2}, \quad \sum_{k=0}^{\infty} {}^2 x^k = \frac{x(1+x)}{(1-x)^3}, \; e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad (a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$
$$X \sim \text{Bern}(p), \quad p_X(0) = 1 - p, \; p_X(1) = p, \; E(X) = p, \; \mathrm{VAR}(X) = p(1-p),$$
$$X \sim \text{Διων}(N, p), \; p_X(k) = \binom{N}{k} p^k (1-p)^{N-k}, \; E(X) = Np, \; \mathrm{VAR}(X) = Np(1-p),$$
$$X \sim \text{Γεωμ}(p), \; p_X(k) = (1-p)^{k-1} p, \; E(X) = 1/p, \; \mathrm{VAR}(X) = (1-p)/p^2, \; P(X \geq m + n | X > n) = P(X \geq m), \; m \geq 1, n \geq 0,$$
$$X \sim \text{Υπερ}(N, k, n), \; p_X(m) = \frac{\binom{k}{m}\binom{N-k}{n-m}}{\binom{N}{n}}, \; E(X) = \frac{nk}{N}, \; \mathrm{VAR}(X) = \frac{nk(N-k)(N-n)}{N^2(N-1)},$$
$$X \sim \text{Poisson}(\lambda), \; p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \; E(X) = \lambda, \; \mathrm{VAR}(X) = \lambda.$$

$$p_{XY}(x, y) = P(X = x, Y = y) \quad \forall x \in S_X, \, \forall y \in S_Y, \quad p_X(x) = \sum_{y \in S_Y} p_{XY}(x, y), \quad p_Y(y) = \sum_{x \in S_X} p_{XY}(x, y),$$
$$E(g(X, Y)) = \sum_{x,y} g(x, y) p_{XY}(x, y), \quad E(aX + bY + c) = aE(X) + bE(Y) + c, \quad E(\sum_{k=1}^K g_k(X, Y)) = \sum_{k=1}^K E(g_k(X, Y)),$$
$$\mathrm{COV}(X, Y) = E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big] = E(XY) - E(X)E(Y), \; \mathrm{VAR}(X + Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\mathrm{COV}(X, Y),$$
$$X, Y \text{ ανεξάρτητες} \Leftrightarrow P(X \in A, Y \in B) = P(X \in A) P(Y \in B) \, \forall A, B \Leftrightarrow p_{XY}(x, y) = p_X(x) p_Y(y) \, \forall x, y,$$
$$X, Y \text{ ανεξάρτητες} \Rightarrow E(g(X)h(Y)) = E(g(X)) E(h(Y)), \quad \mathrm{COV}(X, Y) = 0, \quad p_{X+Y}(m) = \sum_{k=-\infty}^{\infty} p_X(k) p_Y(m - k), \; m \in \mathbb{Z},$$
$$E(\sum_{i=1}^n a_i X_i + b) = \sum_{i=1}^n a_i E(X_i) + b, \quad \mathrm{VAR}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{VAR}(X_i) + 2 \sum_{1 \leq i < j \leq n} \mathrm{COV}(X_i, X_j).$$

$$\int_{-\infty}^{a} f(x)\,dx = \lim_{t\to-\infty}\int_{t}^{a} f(x)\,dx, \int_{a}^{\infty} f(x)\,dx = \lim_{t\to+\infty}\int_{a}^{t} f(x)\,dx, \int_{-\infty}^{\infty} f(x)\,dx = \lim_{t\to-\infty}\int_{t}^{c} f(x)\,dx + \lim_{t\to+\infty}\int_{c}^{t} f(x)\,dx,$$

$$\int_{a}^{b} f(x)\,dx = \lim_{t\to b^-}\int_{a}^{t} f(x)\,dx, \quad \int_{a}^{b} f(x)\,dx = \lim_{t\to a^+}\int_{t}^{b} f(x)\,dx,$$

$$P(X\in B) = \int_{B} f(t)\,dt, \quad P[a\le X\le b] = \int_{a}^{b} f(x)\,dx, \quad F(x) = P(X\le x) = \int_{-\infty}^{x} f(t)\,dt, \quad F'(x) = f(x),$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx, \quad E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx, \quad E(aX+b) = aE(X)+b, \quad E(\sum_{i=1}^{n} a_i g_i(X)) = \sum_{i=1}^{n} a_i E(g_i(X)),$$

$$\sigma^2 = \mathrm{VAR}(X) = E\big[(X-\mu)^2\big] = E(X^2) - (E(X))^2, \quad \mathrm{VAR}(aX+b) = a^2\mathrm{VAR}(X).$$

$$X\sim U[a,b], \quad f(x) = \begin{cases} \frac{1}{b-a}, & \text{για } x\in[a,b], \\ 0, & \text{για } x\notin[a,b], \end{cases} \quad F(x) = \begin{cases} 0, & x<a, \\ \frac{x-a}{b-a}, & a\le x\le b, \\ 1, & x>b, \end{cases} \quad E(X) = \frac{a+b}{2}, \quad \mathrm{VAR}(X) = \frac{(b-a)^2}{12},$$

$$X\sim \text{Εκθ}(\theta), \quad f(x) = \begin{cases} \frac{1}{\theta}e^{-x/\theta}, & \text{για } x\ge 0, \\ 0, & \text{για } x<0, \end{cases} \quad F(x) = \begin{cases} 0, & x\le 0, \\ 1-e^{-\frac{x}{\theta}}, & x\ge 0, \end{cases} \quad E(X) = \theta, \quad \mathrm{VAR}(X) = \theta^2,$$

$$X\sim N(\mu,\sigma^2), \quad f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}, \quad \text{για } x\in\mathbb{R}, \quad E(X) = \mu, \quad \mathrm{VAR}(X) = \sigma^2,$$

$$X\sim N(\mu,\sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1), \quad P(a\le X\le b) = \Phi\Big(\frac{b-\mu}{\sigma}\Big) - \Phi\Big(\frac{a-\mu}{\sigma}\Big),$$

$$Z\sim N(0,1), \quad \phi(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}, \quad \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx, \quad \Phi(x) = 1-\Phi(-x).$$

$$R = \{(x,y) : a\le x\le b,\ \phi_1(x)\le y\le\phi_2(x)\}, \quad \iint_R f(x,y)\,dA = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dy\right) dx,$$

$$R = \{(x,y) : a\le y\le b,\ \phi_1(y)\le x\le\phi_2(y)\}, \quad \iint_R f(x,y)\,dA = \int_a^b \left(\int_{\phi_1(y)}^{\phi_2(y)} f(x,y)\,dx\right) dy,$$

$$\iint_{[a,b]\times[c,d]} f(x,y)\,dA = \int_a^b \left(\int_c^d f(x,y)\,dy\right) dx = \int_c^d \left(\int_a^b f(x,y)\,dx\right) dy,$$

$$P[(X,Y)\in R] = \iint_R f_{XY}(x,y)\,dA, \quad P(a\le X\le b,\ c\le Y\le d) = \iint_{[a,b]\times[c,d]} f_{XY}(x,y)\,dA = \begin{cases} \int_a^b \Big(\int_c^d f_{XY}(x,y)dy\Big)\,dx, \\ \int_c^d \Big(\int_a^b f_{XY}(x,y)dx\Big)\,dy, \end{cases}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx, \quad E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{XY}(x,y)\,dxdy,$$

$$E(aX+bY+c) = aE(X)+bE(Y)+c, \quad E(\sum_{k=1}^{K} g_k(X,Y)) = \sum_{k=1}^{K} E(g_k(X,Y)),$$

$$X,Y \text{ανεξάρτητες} \Leftrightarrow P(X\in A, Y\in B) = P(X\in A)P(Y\in B)\ \forall A,B \Leftrightarrow f_{XY}(x,y) = f_X(x)f_Y(y)\ \forall x,y,$$

$$X,Y \text{ανεξάρτητες} \Rightarrow E(g(X)h(Y)) = E(g(X))E(h(Y)), \quad \mathrm{COV}(X,Y) = 0, \quad f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(t)f_Y(z-t)\,dt,$$

$$E(\sum_{i=1}^{n} a_iX_i + b) = \sum_{i=1}^{n} a_iE(X_i) + b, \quad \mathrm{VAR}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n}\mathrm{VAR}(X_i) + 2\sum_{1\le i<j\le n}\mathrm{COV}(X_i,X_j).$$

$$X\ge 0,\ c>0 \Rightarrow P(X\ge c) \le \frac{E(X)}{c}, \qquad c>0 \Rightarrow P(|X-E(X)|\ge c) \le \frac{\mathrm{VAR}(X)}{c^2},$$

$$X_1, X_2, \ldots \text{ ανεξ. με κοινή κατ.}, \ E(X_i) = \mu,\ \bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i \Rightarrow \begin{cases} \forall\epsilon>0,\ P\big(|\bar{X}_N-\mu|<\epsilon\big) \to 1, \quad \text{καθώς } N\to\infty, \\ P\big(\lim_{N\to\infty}\bar{X}_N = \mu\big) = 1, \end{cases}$$

$$\mathrm{VAR}(X_i) = \sigma^2,\ \bar{S}_N = \frac{1}{\sigma\sqrt{N}}\sum_{i=1}^{N}(X_i-\mu) = \frac{\Big(\sum_{i=1}^{N} X_i\Big) - N\mu}{\sigma\sqrt{N}} \Rightarrow \begin{cases} P\big(a\le\bar{S}_N\le b\big) \to \Phi(b)-\Phi(a), \quad \text{καθώς } N\to\infty, \\ P\big(\bar{S}_N\le a\big) \to \Phi(a), \quad \text{καθώς } N\to\infty, \\ P\big(\bar{S}_N\ge a\big) \to 1-\Phi(a), \quad \text{καθώς } N\to\infty. \end{cases}$$

ΠΙΘΑΝΟΤΗΤΕΣ, 2019-2020
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΤΑΥΡΟΣ ΤΟΥΜΠΗΣ

Λύσεις Τελικής Εξέτασης Ιουνίου Ακ. Έτους 2019-2020

1. **(Αεροπλάνο)** Σε ένα αεροπλάνο επιβαίνουν 40 Βρετανοί, 60 Γάλλοι, και 80 Ρουμάνοι. Καθένας από τους Βρετανούς είναι φορέας του κορωνοϊού με πιθανότητα 20%, ενώ οι αντίστοιχες πιθανότητες για τους Γάλλους και τους Ρουμάνους είναι 10% και 5%.

   (α΄) *(1 μονάδα)* Επιλέγεται ένα άτομο στην τύχη, χωρίς κάποια προτίμηση, και προκύπτει πως είναι φορέας. Ποια είναι η πιθανότητα να είναι Ρουμάνος;

   (β΄) *(1.5 μονάδα)* Επιλέγονται 4 άτομα στην τύχη, χωρίς προτίμηση στο συνδυασμό. Ποια η πιθανότητα να είναι όλοι Ρουμάνοι και επιπλέον ακριβώς 3 από αυτούς να είναι φορείς του κορωνοϊού;

   Λύση:

   (α΄) Έστω B, F, R, τα ενδεχόμενα το άτομο να είναι Βρετανός, Γάλλος, ή Ρουμάνος, αντίστοιχα. Έστω επίσης A το ενδεχόμενο το άτομο να έχει τον κορωναϊό. Έχουμε, με χρήση του Κανόνα του Bayes,

$$P(R|A) = \frac{P(RA)}{P(A)} = \frac{P(A|R)P(R)}{P(A|R)P(R) + P(A|B)P(B) + P(A|F)P(F)}$$
$$= \frac{0.05 \times \frac{8}{18}}{0.05 \times \frac{8}{18} + 0.2 \times \frac{4}{18} + 0.1 \times \frac{6}{18}} = \frac{2}{9}.$$

   (β΄) Έστω B το ενδεχόμενο να είναι και οι 4 Ρουμάνοι, και C το ενδεχόμενο να είναι φορείς ακριβώς οι τρεις. Παρατηρούμε πως

$$P(BC) = P(B)P(C|B) = \frac{\binom{80}{4}}{\binom{180}{4}} \times 4 \times 0.05^3 \times 0.95.$$

   Σχετικά με την $P(B)$, χρησιμοποιήσαμε το γεγονός ότι υπάρχουν 180 άτομα, και από αυτά 80 είναι Ρουμάνοι, οπότε η πιθανότητα $P(B)$ είναι οι συνδυασμοί 4 ατόμων μεταξύ όλων των Ρουμάνων δια τους συνδυασμούς 4 ατόμων μεταξύ όλων των επιβατών. Σχετικά με την $P(C|)$, χρησιμοποιήσαμε το γεγονός ότι αν ξέρουμε ότι είναι όλοι Ρουμάνοι, το πλήθος των φορέων δίνεται από τη διωνυμική κατανομή, αφού κάνουμε ανεξάρτητα πειράματα, καθένα με επιτυχία να είναι το αντίστοιχο άτομο φορέας.

2. **(Υπέρβαση χωρητικότητας)** *(1 μονάδα)* Σε μια εξέταση πιθανοτήτων έχουν λάβει οδηγία να καθίσουν σε μια αίθουσα 22 άτομα. Καθένα από αυτά τα άτομα θα έρθει με πιθανότητα $\frac{3}{4}$. Η αίθουσα χωρά να καθίσουν μόλις 20. Ποια είναι η πιθανότητα p τελικά να υπάρξει τουλάχιστον ένα άτομο που δεν θα βρει θέση να καθίσει σε εκείνη την αίθουσα;

   Λύση: Το πλήθος των ατόμων που θα έρθουν ακολουθεί την διωνυμική κατανομή, με παραμέτρους $N = 22$ και $p = \frac{3}{4}$. Άρα τελικά η πιθανότητα να έρθουν παραπάνω φοιτητές από τις διαθέσιμες θέσεις είναι

$$p = \binom{22}{21}\left(\frac{3}{4}\right)^{21}\left(\frac{1}{4}\right) + \binom{22}{22}\left(\frac{3}{4}\right)^{22} = 22\left(\frac{3}{4}\right)^{21}\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)^{22} = \frac{25}{4}\left(\frac{3}{4}\right)^{21}.$$

3. **(Μετάδοση κορωνοϊού)** *(1 μονάδα)* Σε μια κοινωνική εκδήλωση συμμετέχουν 20 άτομα, εκ των οποίων 5 είναι φορείς του κορωνοϊού. Κάθε ένας από τους φορείς θα μεταδώσει, μέχρι το τέλος της εκδήλωσης, τον κορωνοϊό σε κάθε έναν από τους υπόλοιπους 15 με πιθανότητα 0.1. Ποια είναι η μέση τιμή των ατόμων που θα έχουν γίνει φορείς μέχρι το τέλος της εκδήλωσης; Εννοείται πως αρκεί μια μόνο μετάδοση για να γίνει κάποιος φορέας, και πως οι μεταδόσεις είναι ανεξάρτητες.

   Λύση: Κάθε ένας από τους μη φορείς θα αποφύγει να γίνει φορέας με πιθανότητα 0.9^5, και επομένως θα γίνει φορέας με πιθανότητα $1 - 0.9^5$. Έστω X_i, $i = 1, \ldots, 15$ Τ.Μ. Bernoulli που ισούται με 1 αν ο μη φορέας i γίνει φορέας. Έστω, επίσης, X, το πλήθος των νέων φορέων. Έχουμε

$$E(X) = E\left(\sum_{i=1}^{15} X_i\right) = 15(1 - 0.9^5).$$

4. **(Τοστιέρα-αποστειρωτής)** Ένας διδάσκων πιθανοτήτων αποστειρώνει γραπτά της τελικής εξέτασης στην τοστιέρα του σπιτιού του. Προκειμένου να αποστειρωθεί πλήρως ένα γραπτό, πρέπει να παραμείνει στην τοστιέρα για ένα χρόνο X ο οποίος είναι Τ.Μ. με πυκνότητα

$$f(x) = \begin{cases} C(x-1)e^{-x}, & x \geq 1, \\ 0, & x < 1, \end{cases}$$

όπου το C είναι μια άγνωστη σταθερά.

(α΄) *(1 μονάδα)* Προσδιορίστε την τιμή της σταθεράς C.

(β΄) *(1 μονάδα)* Να δώσετε ένα τύπο για την συνάρτηση κατανομής πιθανότητας $F_X(x)$ της X.

(γ΄) *(1 μονάδα)* Καθώς ο διδάσκων δεν έχει στη διάθεσή του άπειρο χρόνο, κρατά στην τοστιέρα κάθε γραπτό για χρόνο θ ο οποίος εξασφαλίζει ότι το γραπτό θα έχει αποστειρωθεί πλήρως με πιθανότητα 99%. Βρείτε μια εξίσωση που πρέπει να ικανοποιεί το θ. (Μην επιχειρήσετε να την λύσετε ως προς θ, καθώς δεν μπορεί να γραφτεί σε κλειστή μορφή χρησιμοποιώντας απλές συναρτήσεις.)

(δ΄) *(1 μονάδα)* Για τον χρόνο θ του προηγούμενου σκέλους, ποια ακριβώς είναι η κατανομή των γραπτών που ΔΕΝ έχουν αποστειρωθεί πλήρως, αν υποθέσουμε ότι υπάρχουν 400 γραπτά και οι χρόνοι αποστείρωσης είναι ανεξάρτητοι μεταξύ τους; Ποια είναι επακριβώς η πιθανότητα να μην έχουν αποστειρωθεί ακριβώς 10 γραπτά; Μπορείτε να γράψετε μια προσέγγιση για αυτή την πιθανότητα;

Λύση:

(α΄) Θα πρέπει το ολοκλήρωμα της πυκνότητας σε όλο το $\mathbb{R}$ να ισούται με τη μονάδα. Το ολοκλήρωμα όμως αυτό ισούται με

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \int_1^{\infty} C(x-1)e^{-x}\,dx = C\int_1^{\infty}(x-1)(-e^{-x})'\,dx \\ &= \left[-C(x-1)e^{-x}\right]_1^{\infty} + C\int_1^{\infty} e^{-x}\,dx = 0 + C\left[e^{-x}\right]_{\infty}^{1} = Ce^{-1} = 1 \Rightarrow C = e. \end{aligned}$$

(β΄) Γενικά ισχύει $F(x) = \int_{-\infty}^{x} f(t)\,dt$, επομένως η $F(x) = 0$ για $x < 1$, ενώ για $x \geq 1$ έχουμε

$$\begin{aligned} F(x) &= \int_1^x e(t-1)e^{-t}\,dt = -\int_1^x e(t-1)(e^{-t})'\,dt = \left[-e(t-1)e^{-t}\right]_1^x + e\int_1^x e^{-t}\,dt \\ &= -e^{1-x}(x-1) - e\int_1^x (e^{-t})'\,dt = -(x-1)e^{1-x} - e(e^{-x} - e^{-1}) \\ &= -(x-1)e^{-(x-1)} - e^{-(x-1)} + 1 = (-x+1-1)e^{-(x-1)} + 1 = 1 - xe^{-(x-1)}, \end{aligned}$$

άρα, συγκεντρωτικά,

$$F(x) = \begin{cases} 1 - xe^{(x-1)}, & x \geq 1, \\ 0, & x < 1. \end{cases}$$

(γ΄) Πρέπει

$$1 - \theta e^{-(\theta-1)} = 0.99 \Leftrightarrow \theta e^{-(\theta-1)} = 0.01.$$

(δ΄) Το πλήθος των γραπτών που δεν αποστειρώνεται ακολουθεί τη διωνυμική κατανομή με παραμέτρους $N = 400$ και $p = 0.01$, διότι το κάθε γραπτό από τα 400 δεν αποστειρώνεται ανεξάρτητα από τα άλλα γραπτά με πιθανότητα 0.99. Επομένως, η πιθανότητα να μην αποστειρωθούν ακριβώς 10 γραπτά είναι

$$p = \binom{400}{10} 0.01^{10} 0.99^{390}.$$

Προσεγγιστικά, επειδή το N είναι πολύ μεγάλο και το p πολύ μικρό, ή άνω πιθανότητα δίνεται προσεγγιστικά και από την κατανομή Poisson με παράμετρο $\lambda = 400 \times 0.01 = 4$:

$$p \simeq e^{-4}\frac{4^{10}}{10!}.$$

5. **(Δηλώσεις συμμετοχής)** *(1.5 μονάδα)* Σε ένα μάθημα πιθανοτήτων έχουν δηλώσει συμμετοχή 600 άτομα. Καθένας όμως από αυτούς θα προσέλθει στο μάθημα με πιθανότητα $\frac{2}{3}$, ανεξάρτητα από τους υπόλοιπους. Να υπολογίσετε προσεγγιστικά την πιθανότητα να έρθουν, τελικά, πάνω από 420 άτομα.

Λύση: Έστω Τ.Μ. Bernoulli X_i που είναι ίσες με 1 αν ο φοιτητής i προσέλθει και 0 αλλιώς. Παρατηρήστε πως η διασπορά τους είναι $\mathrm{VAR}(X_i) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$, επομένως έχουν τυπική απόκλιση $\sigma = \frac{\sqrt{2}}{3}$. Έχουμε, με χρήση του Κ.Ο.Θ.,

$$P\left(\sum_{i=1}^{600} X_i > 420\right) = P\left(\frac{\sum_{i1=}^{600} X_i - 400}{\sqrt{600}\frac{\sqrt{2}}{3}} > \frac{420-400}{\sqrt{600}\frac{\sqrt{2}}{3}}\right) \simeq 1 - \Phi\left(\frac{60}{10\sqrt{12}}\right) = 1 - \Phi\left(\sqrt{3}\right).$$

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ, ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΙΘΑΝΟΤΗΤΕΣ, ΣΤΑΥΡΟΣ ΤΟΥΜΠΗΣ
ΤΕΛΙΚΗ ΕΞΕΤΑΣΗ, ΣΕΠΤΕΜΒΡΙΟΣ 2020
ΟΝΟΜΑ ΦΟΙΤΗΤΗ: . . . .

Οδηγίες

1. **Συμπληρώστε το όνομά σας άνω, και παραδώστε το παρόν με τις λύσεις**.
2. Διάρκεια εξέτασης: 100 ΛΕΠΤΑ
3. Απαγορεύεται η αναχώρηση από την αίθουσα πριν την συμπλήρωση 30λέπτου.
4. Απαγορεύεται η χρήση υπολογιστή χειρός. Απαγορεύεται η χρήση κινητού, και ως υπολογιστή χειρός.
5. Οι λύσεις πρέπει να γραφούν αποκλειστικά στην παρεχόμενη κόλλα.
6. Μπορείτε να χρησιμοποιείτε μολύβι ή/και στυλό οποιουδήποτε χρώματος εκτός από κόκκινο.
7. Οι λύσεις πρέπει να είναι το κατά δυνατόν αναλυτικές. **Πρέπει να φαίνονται όλα τα ενδιάμεσα βήματα στους υπολογισμούς. Τοποθετήστε τα τελικά αποτελέσματα εντός πλαισίου.**
8. Ξεκινήστε από αυτές τις ασκήσεις που ξέρετε και/ή δίνουν πολλές μονάδες.
9. Προχωρήστε κάθε άσκηση όσο μπορείτε! Θα δοθούν μονάδες για ασκήσεις λυμένες εν μέρει.
10. Στις τελικές απαντήσεις αρκεί να δοθούν μαθηματικές εκφράσεις, δεν χρειάζεται να κάνετε πράξεις. Μπορείτε να χρησιμοποιήσετε αιτιολογημένες προσεγγίσεις.
11. Σύμφωνα με τον κανονισμό εξετάσεων, σε περίπτωση που διαπιστωθεί (κατά τη διάρκεια της εξέτασης είτε της διόρθωσης) αντιγραφή είτε απόπειρα αντιγραφής, θα ενημερωθούν τα αρμόδια όργανα του ιδρύματος.

ΘΕΜΑΤΑ

1. **(Δύο φίλοι)** *(1.5 μονάδα)* Δύο φίλοι συναντιούνται σε μια καφετέρια. Ο ένας έχει τον κορωνοϊό, ενώ ο άλλος όχι. Καθένας από τους δύο φίλους θα φορέσει μάσκα με πιθανότητα $p = 0.5$, ανεξάρτητα από τον άλλο. Αν δεν φορέσει κανένας μάσκα, ο υγιής θα κολλήσει με πιθανότητα 0.5. Αν φοράει μόνο ο ένας, ο υγιής θα κολλήσει με πιθανότητα 0.25, και αν φοράνε και οι δύο, ο υγιής θα κολλήσει με πιθανότητα 0.1. Με δεδομένο ότι τελικά ο υγιής κόλλησε, ποια η πιθανότητα να φορούσαν και οι δύο μάσκες;

2. **(Άλλοι δύο φίλοι)** Σε ένα άλλο τραπέζι της καφετέριας, συνομιλούν δύο φίλοι, ο A και ο B, εκ των οποίων ο A είναι υγιής, ενώ ο B έχει κορωνοϊό. Όποτε ο B φτερνίζεται, κολλά τον A με κορωνοϊό με πιθανότητα $p_1 = 0.05$. Τα διαδοχικά φτερνίσματα είναι ανεξάρτητα μεταξύ τους.
   (α΄) *(0.5 μονάδα)* Ποια κατανομή ακολουθεί το πλήθος των φορών που πρέπει να φτερνιστεί ο B για να κολλήσει τον A, και με ποια παράμετρο ή παραμέτρους;
   (β΄) *(1 μονάδα)* Πόσες φορές πρέπει να φτερνιστεί ο B ώστε ο A να έχει κολλήσει, σε κάποιο από τα φτερνίσματα, τον κορωνοϊό, με πιθανότητα τουλάχιστον 0.8;

3. **(Δέκα φίλοι)** Σε ένα τρίτο τραπέζι της καφετέριας συναντιούνται 10 φίλοι. Καθένας από αυτούς είναι **νυν φορέας** του κορωνοϊού, δηλαδή ήταν φορέας στην αρχή της συνάντησης, με πιθανότητα $p = 0.3$ ανεξάρτητα από τους άλλους. Κάθε νυν φορέας θα μεταδώσει τον ιό σε κάθε έναν που δεν έχει τον ιό με πιθανότητα $q = \frac{1}{4}$, και θα τον κάνει **νέο φορέα**, ανεξάρτητα από το τι θα συμβεί με τα άλλα ζεύγη ατόμων. Κάποιος γίνεται νέος φορέας ακόμα και αν τον κολλήσει μόνο ένας νυν φορέας. Οι νέοι φορείς δεν κολλούν άλλους.
   (α΄) *(1 μονάδα)* Έστω X το πλήθος των νυν φορέων. Ποια είναι η συνάρτηση μάζας πιθανότητας της Τ.Μ. X; Γράψτε αναλυτικό τύπο.

(β′) *(1 μονάδα)* Αν στην αρχή της συνάντησης είχαν τρεις φίλοι τον κορωνοϊό, δηλαδή $X = 3$, ποια είναι η μάζα της Τ.Μ. Y των νέων φορέων;

4. **(Εμβόλια)** Δύο φαρμακευτικές εταιρίες ετοιμάζουν εμβόλια για τον κορονοϊό. Το πρώτο εξ αυτών θα είναι έτοιμο μετά από ένα τυχαίο χρόνο X, μετρούμενο σε έτη, που ακολουθεί την πυκνότητα

$$f(x) = \begin{cases} k_1 e^{-x}, & 0 \leq x \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

Το δεύτερο εξ αυτών θα είναι έτοιμο μετά από τυχαίο χρόνο Y, επίσης μετρούμενο σε έτη, που ακολουθεί την πυκνότητα πιθανότητας

$$g(y) = \begin{cases} k_2 y(1-y), & 0 \leq y \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

Οι X, Y είναι ανεξάρτητες μεταξύ τους.

(α′) *(1 μονάδα)* Προσδιορίστε την τιμή της k_1.

(β′) *(1 μονάδα)* Προσδιορίστε την τιμή της k_2.

(γ′) *(1 μονάδα)* Ποια είναι η πιθανότητα να χρειαστεί να περιμένουμε μισό τουλάχιστον χρόνο για να γίνει διαθέσιμο κάποιο εμβόλιο, είτε το ένα είτε το άλλο;

5. **(Δοκιμές εμβολίων)** Μια φαρμακευτική εταιρία δοκιμάζει την αποτελεσματικότητα ενός εμβολίου, ως εξής. Η εταιρεία εμβολιάζει $N = 10^4$ άτομα και ακολούθως τα αφήνει να συνεχίσουν τις καθημερινές τους δραστηριότητες. Μετά από δύο μήνες, τα υποβάλει όλα σε εξέταση για να διαπιστώσει ποιοι είναι φορείς. Η εταιρεία γνωρίζει ότι αν το εμβόλιο είναι αποτελεσματικό, τότε κάθε άτομο θα είναι φορέας με μια μικρή πιθανότητα $p = 0.1$. Αντίθετα, αν τελικά το εμβόλιο δεν είναι αποτελεσματικό, τότε κάθε άτομο θα είναι φορέας με πιθανότητα $p = 0.8$.

(α′) *(0.5 μονάδα)* Υποθέτοντας ότι το εμβόλιο είναι ΜΗ αποτελεσματικό, κατά μέσο όρο πόσα από τα N άτομα θα νοσήσουν;

(β′) *(1.5 μονάδα)* Υποθέτοντας ότι το εμβόλιο είναι αποτελεσματικό, ποια είναι, κατά προσέγγιση, η πιθανότητα οι φορείς να είναι, εν τέλει, άνω των 1100 ατόμων;

ΤΥΠΟΛΟΓΙΟ ΠΙΘΑΝΟΤΗΤΩΝ

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad (A \cup B)' = A' \cap B', \quad (A \cap B)' = A' \cup B',$$
$$P(A) \geq 0 \, \forall A \in \mathcal{F}, \quad P(\Omega) = 1, \quad A_i \cap A_j = \emptyset \Rightarrow P({}_1 \cup {}_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots,$$
$$P(A') = 1 - P(A), \quad P(\emptyset) = 0, \quad P(A) \leq 1, \quad A \subseteq B \Rightarrow P(A) \leq P(B), \quad P(A \cap B') = P(A) - P(A \cap B),$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \; P(A \cup B) \leq P(A) + P(B), \; P(A \cup B \cup C) = P(A) + P(A' \cap B) + P(A' \cap B' \cap C),$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C), \; P(A) = P(A \cap B) + P(A \cap B').$$

Διατάξεις k αντικ. από N: $\frac{N!}{(N-k)!}$, Συνδυασμοί k αντικ. από N: $\binom{N}{k} = \frac{N!}{k!(N-k)!}$,
Επαναληπτικές διατάξεις μήκους k από αντικ.: N^k, Επαναληπτικοί συνδυασμοί k αντικ. από N: $\binom{+k-1}{k} = \binom{+k-1}{-1}$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(A_1 A_2 \dots A_n) = P(A_1) P(A_2|A_1) P(A_3|A_2 A_1) \dots P(A_n|A_1 A_2 \dots A_{n-1}),$$
$$P() = P(A|B)P(B) + P(A|B')P(B'), \quad P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B')P(B')},$$
$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N) \quad (\{B_i\} \text{ διαμέριση}),$$
$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N)} \quad (\{B_i\} \text{ διαμέριση}).$$
$$A, B \text{ ανεξάρτητα} \Leftrightarrow P(AB) = P(A)P(B) \Leftrightarrow P(B) = P(B|A) \Leftrightarrow P(A) = P(A|B).$$

$$p_X(x) = P(X = x) = P(\{\omega : X(\omega) = x\}) \, \forall x \in S_X, \quad F_X(x) = P(X \leq x), \quad \lim_{t \to x^-} F(t) = P(X < x), \quad E(X) = \sum_{x \in S_X} x p_X(x),$$
$$E(g(X)) = \sum_{x \in S_X} g(x) p_X(x), \quad E(aX + b) = aE(X) + b, \quad E(\sum_{i=1}^{n} a_i g_i(X)) = \sum_{i=1}^{n} a_i E(g_i(X)),$$
$$\sigma^2 = \mathrm{VAR}(X) = E\big[(X - \mu)^2\big] = E(X^2) - (E(X))^2, \quad \mathrm{VAR}(aX + b) = a^2 \mathrm{VAR}(X).$$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad \sum_{k=0}^{\infty} k x^k = \frac{x}{(1-x)^2}, \quad \sum_{k=0}^{\infty} {}^2 x^k = \frac{x(1+x)}{(1-x)^3}, \; e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad (a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$
$$X \sim \text{Bern}(p), \quad p_X(0) = 1 - p, \; p_X(1) = p, \; E(X) = p, \; \mathrm{VAR}(X) = p(1-p),$$
$$X \sim \text{Διων}(N, p), \; p_X(k) = \binom{N}{k} p^k (1-p)^{N-k}, \; E(X) = Np, \; \mathrm{VAR}(X) = Np(1-p),$$
$$X \sim \text{Γεωμ}(p), \; p_X(k) = (1-p)^{k-1} p, \; E(X) = 1/p, \; \mathrm{VAR}(X) = (1-p)/p^2, \; P(X \geq m + n | X > n) = P(X \geq m), \; m \geq 1, n \geq 0,$$
$$X \sim \text{Υπερ}(N, k, n), \; p_X(m) = \frac{\binom{k}{m}\binom{N-k}{n-m}}{\binom{N}{n}}, \; E(X) = \frac{nk}{N}, \; \mathrm{VAR}(X) = \frac{nk(N-k)(N-n)}{N^2(N-1)},$$
$$X \sim \text{Poisson}(\lambda), \; p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \; E(X) = \lambda, \; \mathrm{VAR}(X) = \lambda.$$

$$p_{XY}(x, y) = P(X = x, Y = y) \quad \forall x \in S_X, \, \forall y \in S_Y, \quad p_X(x) = \sum_{y \in S_Y} p_{XY}(x, y), \quad p_Y(y) = \sum_{x \in S_X} p_{XY}(x, y),$$
$$E(g(X, Y)) = \sum_{x, y} g(x, y) p_{XY}(x, y), \quad E(aX + bY + c) = aE(X) + bE(Y) + c, \quad E(\sum_{k=1}^{K} g_k(X, Y)) = \sum_{k=1}^{K} E(g_k(X, Y)),$$
$$\mathrm{COV}(X, Y) = E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big] = E(XY) - E(X)E(Y), \; \mathrm{VAR}(X + Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\mathrm{COV}(X, Y),$$
$$X, Y \text{ ανεξάρτητες} \Leftrightarrow P(X \in A, Y \in B) = P(X \in A) P(Y \in B) \, \forall A, B \Leftrightarrow p_{XY}(x, y) = p_X(x) p_Y(y) \, \forall x, y,$$
$$X, Y \text{ ανεξάρτητες} \Rightarrow E(g(X)h(Y)) = E(g(X))E(h(Y)), \quad \mathrm{COV}(X, Y) = 0, \quad p_{X+Y}(m) = \sum_{k=-\infty}^{\infty} p_X(k) p_Y(m-k), \; m \in \mathbb{Z},$$
$$E(\sum_{i=1}^{n} a_i X_i + b) = \sum_{i=1}^{n} a_i E(X_i) + b, \quad \mathrm{VAR}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{VAR}(X_i) + 2 \sum_{1 \leq i < j \leq n} \mathrm{COV}(X_i, X_j).$$

$$\int_{-\infty}^{a} f(x)\,dx = \lim_{t\to-\infty}\int_t^a f(x)\,dx,\ \int_a^{\infty} f(x)\,dx = \lim_{t\to+\infty}\int_a^t f(x)\,dx,\ \int_{-\infty}^{\infty} f(x)\,dx = \lim_{t\to-\infty}\int_t^c f(x)\,dx + \lim_{t\to+\infty}\int_c^t f(x)\,dx,$$

$$\int_a^b f(x)\,dx = \lim_{t\to b^-}\int_a^t f(x)\,dx, \quad \int_a^b f(x)\,dx = \lim_{t\to a^+}\int_t^b f(x)\,dx,$$

$$P(X\in B) = \int_B f(t)\,dt, \quad P[a\le X\le b] = \int_a^b f(x)\,dx, \quad F(x) = P(X\le x) = \int_{-\infty}^x f(t)\,dt, \quad F'(x) = f(x),$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx, \quad E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx, \quad E(aX+b) = aE(X)+b, \quad E(\sum_{i=1}^n a_i g_i(X)) = \sum_{i=1}^n a_i E(g_i(X)),$$

$$\sigma^2 = \mathrm{VAR}(X) = E\big[(X-\mu)^2\big] = E(X^2) - (E(X))^2, \quad \mathrm{VAR}(aX+b) = a^2\mathrm{VAR}(X).$$

$$X\sim U[a,b], \quad f(x) = \begin{cases}\frac{1}{b-a}, & \text{για } x\in[a,b],\\ 0, & \text{για } x\notin[a,b],\end{cases} \quad F(x) = \begin{cases}0, & x<a,\\ \frac{x-a}{b-a}, & a\le x\le b,\\ 1, & x>b,\end{cases} \quad E(X) = \frac{a+b}{2}, \quad \mathrm{VAR}(X) = \frac{(b-a)^2}{12},$$

$$X\sim \text{Εκθ}(\theta), \quad f(x) = \begin{cases}\frac{1}{\theta}e^{-x/\theta}, & \text{για } x\ge 0,\\ 0, & \text{για } x<0,\end{cases} \quad F(x) = \begin{cases}0, & x\le 0,\\ 1-e^{-\frac{x}{\theta}}, & x\ge 0,\end{cases} \quad E(X) = \theta, \quad \mathrm{VAR}(X) = \theta^2,$$

$$X\sim N(\mu,\sigma^2), \quad f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}, \quad \text{για } x\in\mathbb{R}, \quad E(X) = \mu, \quad \mathrm{VAR}(X) = \sigma^2,$$

$$X\sim N(\mu,\sigma^2) \Rightarrow \frac{X-\mu}{\sigma}\sim N(0,1), \quad P(a\le X\le b) = \Phi\Big(\frac{b-\mu}{\sigma}\Big) - \Phi\Big(\frac{a-\mu}{\sigma}\Big),$$

$$Z\sim N(0,1), \quad \phi(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}, \quad \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx, \quad \Phi(x) = 1-\Phi(-x).$$

$$R = \{(x,y) : a\le x\le b,\ \phi_1(x)\le y\le\phi_2(x)\}, \quad \iint_R f(x,y)\,dA = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dy\right)dx,$$

$$R = \{(x,y) : a\le y\le b,\ \phi_1(y)\le x\le\phi_2(y)\}, \quad \iint_R f(x,y)\,dA = \int_a^b \left(\int_{\phi_1(y)}^{\phi_2(y)} f(x,y)\,dx\right)dy,$$

$$\iint_{[a,b]\times[c,d]} f(x,y)\,dA = \int_a^b\left(\int_c^d f(x,y)\,dy\right)dx = \int_c^d\left(\int_a^b f(x,y)\,dx\right)dy,$$

$$P[(X,Y)\in R] = \iint_R f_{XY}(x,y)\,dA, \quad P(a\le X\le b,\ c\le Y\le d) = \iint_{[a,b]\times[c,d]} f_{XY}(x,y)\,dA = \begin{cases}\int_a^b\Big(\int_c^d f_{XY}(x,y)dy\Big)dx,\\ \int_c^d\Big(\int_a^b f_{XY}(x,y)dx\Big)dy,\end{cases}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx, \quad E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{XY}(x,y)\,dxdy,$$

$$E(aX+bY+c) = aE(X)+bE(Y)+c, \quad E(\sum_{k=1}^K g_k(X,Y)) = \sum_{k=1}^K E(g_k(X,Y)),$$

$$X,Y\,\text{ανεξάρτητες} \Leftrightarrow P(X\in A, Y\in B) = P(X\in A)P(Y\in B)\ \forall A,B \Leftrightarrow f_{XY}(x,y) = f_X(x)f_Y(y)\ \forall x,y,$$

$$X,Y\,\text{ανεξάρτητες} \Rightarrow E(g(X)h(Y)) = E(g(X))E(h(Y)), \quad \mathrm{COV}(X,Y) = 0, \quad f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(t)f_Y(z-t)\,dt,$$

$$E(\sum_{i=1}^n a_iX_i+b) = \sum_{i=1}^n a_iE(X_i)+b, \quad \mathrm{VAR}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{VAR}(X_i) + 2\sum_{1\le i<j\le n}\mathrm{COV}(X_i,X_j).$$

$$X\ge 0,\ c>0 \Rightarrow P(X\ge c)\le\frac{E(X)}{c}, \qquad c>0\Rightarrow P(|X-E(X)|\ge c)\le\frac{\mathrm{VAR}(X)}{c^2},$$

$$X_1, X_2, \ldots\ \text{ανεξ. με κοινή κατ.},\ E(X_i) = \mu,\ \bar{X}_N = \frac{1}{N}\sum_{i=1}^N X_i \Rightarrow \begin{cases}\forall\epsilon>0,\ P\big(|\bar{X}_N-\mu|<\epsilon\big)\to 1, \quad \text{καθώς } N\to\infty,\\ P\big(\lim_{N\to\infty}\bar{X}_N = \mu\big) = 1,\end{cases}$$

$$\mathrm{VAR}(X_i) = \sigma^2,\ \bar{S}_N = \frac{1}{\sigma\sqrt{N}}\sum_{i=1}^N (X_i-\mu) = \frac{\Big(\sum_{i=1}^N X_i\Big)-N\mu}{\sigma\sqrt{N}} \Rightarrow \begin{cases}P\big(a\le\bar{S}_N\le b\big)\to\Phi(b)-\Phi(a), \quad \text{καθώς } N\to\infty,\\ P\big(\bar{S}_N\le a\big)\to\Phi(a), \quad \text{καθώς } N\to\infty,\\ P\big(\bar{S}_N\ge a\big)\to 1-\Phi(a), \quad \text{καθώς } N\to\infty.\end{cases}$$

ΠΙΘΑΝΟΤΗΤΕΣ, 2019-2020
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΤΑΥΡΟΣ ΤΟΥΜΠΗΣ

Λύσεις Τελικής Εξέτασης Σεπτεμβρίου Ακ. Έτους 2019-2020

1. **(Δύο φίλοι)** *(1.5 μονάδα)* Δύο φίλοι συναντιούνται σε μια καφετέρια. Ο ένας έχει τον κορωνοϊό, ενώ ο άλλος όχι. Καθένας από τους δύο φίλους θα φορέσει μάσκα με πιθανότητα $p = 0.5$, ανεξάρτητα από τον άλλο. Αν δεν φορέσει κανένας μάσκα, ο υγιής θα κολλήσει με πιθανότητα 0.5. Αν φοράει μόνο ο ένας, ο υγιής θα κολλήσει με πιθανότητα 0.25, και αν φοράνε και οι δύο, ο υγιής θα κολλήσει με πιθανότητα 0.1. Με δεδομένο ότι τελικά ο υγιής κόλλησε, ποια η πιθανότητα να φορούσαν και οι δύο μάσκες;

   Λύση: Έστω B_0, B_1, και B_2 τα ενδεχόμενα να φορούσαν οι φίλοι καμία, μια, και δύο μάσκες, αντίστοιχα. Έχουμε

   $$P(B_0) = 0.25, \quad P(B_1) = 0.5, \quad P(B_2) = 0.1.$$

   Έστω, επίσης, το ενδεχόμενο K να κολλήσει ο υγιής φίλος. Έχουμε

   $$P(B_2|K) = \frac{P(K|B_2)P(B_2)}{P(K|B_0)P(B_0) + P(K|B_1)P(B_1) + P(K|B_2)P(B_2)}$$
   $$= \frac{0.1 \times 0.25}{0.5 \times 0.25 + 0.25 \times 0.5 + 0.1 \times 0.25} = \frac{1}{11}.$$

2. **(Άλλοι δύο φίλοι)** Σε ένα άλλο τραπέζι της καφετέριας, συνομιλούν δύο φίλοι, ο A και ο B, εκ των οποίων ο A είναι υγιής, ενώ ο B έχει κορωνοϊό. Όποτε ο B φτερνίζεται, κολλά τον A με κορωνοϊό με πιθανότητα $p_1 = 0.05$. Τα διαδοχικά φτερνίσματα είναι ανεξάρτητα μεταξύ τους.

   (α΄) *(0.5 μονάδα)* Ποια κατανομή ακολουθεί το πλήθος των φορών που πρέπει να φτερνιστεί ο B για να κολλήσει τον A, και με ποια παράμετρο ή παραμέτρους;

   (β΄) *(1 μονάδα)* Πόσες φορές πρέπει να φτερνιστεί ο B ώστε ο A να έχει κολλήσει, σε κάποιο από τα φτερνίσματα, τον κορωνοϊό, με πιθανότητα τουλάχιστον 0.8;

   Λύση:

   (α΄) Τα φτερνίσματα είναι ανεξάρτητα και όμοια, επομένως αν θεωρήσουμε επιτυχία να κολλήσει ο A και X είναι το πλήθος μέχρι το πρώτο επιτυχημένο φτέρνισμα, η X ακολουθεί τη γεωμετρική κατανομή, με παράμετρο $p = 0.05$.

   (β΄) Κατά τα γνωστά από τη θεωρία για την γεωμετρική κατανομή, $P(X > k) = (1-p)^k$. Επομένως,

   $$P(X \le k) \ge 0.8 \Leftrightarrow 1 - P(X > k) \ge 0.8 \Leftrightarrow P(X < k) \le 0.2$$
   $$\Leftrightarrow (1-p)^k \le 0.2 \Leftrightarrow k \log(1-p) \le \log 0.2 \Leftrightarrow k \ge \frac{\log 0.2}{\log(1-p)}.$$

3. **(Δέκα φίλοι)** Σε ένα τρίτο τραπέζι της καφετέριας συναντιούνται 10 φίλοι. Καθένας από αυτούς είναι **νυν φορέας** του κορωνοϊού, δηλαδή ήταν φορέας στην αρχή της συνάντησης, με πιθανότητα $p = 0.3$ ανεξάρτητα από τους άλλους. Κάθε νυν φορέας θα μεταδώσει τον ιό σε κάθε έναν που δεν έχει τον ιό με πιθανότητα $q = \frac{1}{4}$, και θα τον κάνει **νέο φορέα**, ανεξάρτητα από το τι θα συμβεί με τα άλλα ζεύγη ατόμων. Κάποιος γίνεται νέος φορέας ακόμα και αν τον κολλήσει μόνο ένας νυν φορέας. Οι νέοι φορείς δεν κολλούν άλλους.

   (α΄) *(1 μονάδα)* Έστω X το πλήθος των νυν φορέων. Ποια είναι η συνάρτηση μάζας πιθανότητας της Τ.Μ. X; Γράψτε αναλυτικό τύπο.

   (β΄) *(1 μονάδα)* Αν στην αρχή της συνάντησης είχαν τρεις φίλοι τον κορωνοϊό, δηλαδή $X = 3$, ποια είναι η μάζα της Τ.Μ. Y των νέων φορέων;

   Λύση:

   (α΄) Κάθε ένας από τους 10 φίλους θα έχει τον ιό ανεξάρτητα από τους άλλους, επομένως η X ακολουθεί την διωνυμική κατανομή με παραμέτρους $N = 10$ και $p = 0.3$. Επομένως,

   $$P(X = k) = \binom{10}{k} p^k (1-p)^{10-k}.$$

(β΄) Υπάρχουν $N = 7$ φίλοι που μπορεί να κολλήσουν τον κορωνοϊό, και καθένας από αυτούς θα κολλήσει εφόσον τον κολλήσει ένας από όσους έχουν τον κορωνοϊό, κάτι που γίνεται με πιθανότητα $q = 1 - \left(\frac{3}{4}\right)^3$. Επομένως, η ζητούμενη μάζα είναι η

$$P(Y = k) = \binom{7}{k} q^k (1 - q)^{1-k}.$$

4. **(Εμβόλια)** Δύο φαρμακευτικές εταιρίες ετοιμάζουν εμβόλια για τον κορονοϊό. Το πρώτο εξ αυτών θα είναι έτοιμο μετά από ένα τυχαίο χρόνο X, μετρούμενο σε έτη, που ακολουθεί την πυκνότητα

$$f(x) = \begin{cases} k_1 e^{-x}, & 0 \leq x \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

Το δεύτερο εξ αυτών θα είναι έτοιμο μετά από τυχαίο χρόνο Y, επίσης μετρούμενο σε έτη, που ακολουθεί την πυκνότητα πιθανότητας

$$g(y) = \begin{cases} k_2 y(1 - y), & 0 \leq y \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

Οι X, Y είναι ανεξάρτητες μεταξύ τους.

(α΄) *(1 μονάδα)* Προσδιορίστε την τιμή της k_1.

(β΄) *(1 μονάδα)* Προσδιορίστε την τιμή της k_2.

(γ΄) *(1 μονάδα)* Ποια είναι η πιθανότητα να χρειαστεί να περιμένουμε μισό τουλάχιστον χρόνο για να γίνει διαθέσιμο κάποιο εμβόλιο, είτε το ένα είτε το άλλο;

Λύση:

(α΄) Κατά τα γνωστά από τη θεωρία,

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \Leftrightarrow k_1 \int_0^1 e^{-x}\, dx = 1 \Leftrightarrow k_1 \left[-e^{-x}\right]_0^1 = 1 \Leftrightarrow k_1 = \frac{1}{1 - e^{-1}}.$$

(β΄) Ομοίως,

$$\int_{-\infty}^{\infty} g(y)\, dy = 1 \Leftrightarrow k_2 \int_0^1 y(1 - y)\, dy = 1 \Leftrightarrow k_2 \left[\right]_0^1 = 1 \Leftrightarrow k_2 \left[\frac{1}{2} - \frac{1}{3}\right] = 1 \Leftrightarrow k_2 = 6.$$

(γ΄) Θα περιμένουμε μισό τουλάχιστον χρόνο, αν και τα δύο εμβόλιο χρειαστούν πάνω από μισό χρόνο. Λόγω ανεξαρτησίας, αυτό θα γίνει με πιθανότητα $P\left(X > \frac{1}{2}\right) P\left(Y > \frac{1}{2}\right)$. Όμως

$$P\left(X \geq \frac{1}{2}\right) = k_1 \int_{\frac{1}{2}}^1 e^{-x}\, dx = k_1 \left[-e^{-x}\right]_{\frac{1}{2}}^1 = \frac{1 - e^{-\frac{1}{2}}}{1 - e^{-1}},$$

και

$$P\left(Y \geq \frac{1}{2}\right) = k_2 \int_{\frac{1}{2}}^1 y(1 - y)\, dy = k_2 \left[\frac{y^2}{2} - \frac{y^3}{3}\right]_{\frac{1}{2}}^1 = 6\left[\frac{1}{2} - \frac{1}{3} - \frac{1}{8} + \frac{1}{24}\right] = \frac{1}{2}.$$

Άρα τελικά η ζητούμενη πιθανότητα είναι

$$\frac{1 - e^{-\frac{1}{2}}}{2(1 - e^{-1})}.$$

5. **(Δοκιμές εμβολίων)** Μια φαρμακευτική εταιρία δοκιμάζει την αποτελεσματικότητα ενός εμβολίου, ως εξής. Η εταιρεία εμβολιάζει $N = 10^4$ άτομα και ακολούθως τα αφήνει να συνεχίσουν τις καθημερινές τους δραστηριότητες. Μετά από δύο μήνες, τα υποβάλει όλα σε εξέταση για να διαπιστώσει ποιοι είναι φορείς. Η εταιρεία γνωρίζει ότι αν το εμβόλιο είναι αποτελεσματικό, τότε κάθε άτομο θα είναι φορέας με μια μικρή πιθανότητα $p = 0.1$. Αντίθετα, αν τελικά το εμβόλιο δεν είναι αποτελεσματικό, τότε κάθε άτομο θα είναι φορέας με πιθανότητα $p = 0.8$.

(α΄) *(0.5 μονάδα)* Υποθέτοντας ότι το εμβόλιο είναι ΜΗ αποτελεσματικό, κατά μέσο όρο πόσα από τα N άτομα θα νοσήσουν;

(β΄) *(1.5 μονάδα)* Υποθέτοντας ότι το εμβόλιο είναι αποτελεσματικό, ποια είναι, κατά προσέγγιση, η πιθανότητα οι φορείς να είναι, εν τέλει, άνω των 1100 ατόμων;

Λύση: Έστω Bernoulli Τ.Μ. $X_i = 1$ αν νοσήσει το άτομο i, και $X_i = 0$ αν δεν νοσήσει, με $i = 1, \ldots, 10^4$. Στην περίπτωση που το εμβόλιο είναι αποτελεσματικό, η τυπική απόκλιση αυτών των Τ.Μ. είναι

$$\sigma = \sqrt{\frac{1}{10} \times \frac{9}{10}} = \frac{3}{10}.$$

Στην περίπτωση που το εμβόλιο δεν είναι αποτελεσματικό, η τυπική απόκλιση αυτών των Τ.Μ. είναι

$$\sigma = \sqrt{\frac{1}{5} \times \frac{4}{5}} = \frac{2}{5}.$$

(α΄) Έχουμε

$$E\left(\sum_{i=1}^{N} X_i\right) = NE(X_i) = 8000.$$

(β΄) Χρησιμοποιώντας το Κ.Ο.Θ.:

$$\begin{aligned} P(X > 1100) = P\left(\sum_{i=1}^{N} X_i > 1100\right) &= P\left(\frac{\sum_{i=1}^{N} X_i - Np}{\sigma\sqrt{N}} > \frac{1100 - Np}{\sigma\sqrt{N}}\right) \\ &= 1 - \Phi\left(\frac{1100 - Np}{\sigma\sqrt{N}}\right) = 1 - \Phi\left(\frac{(1100 - 1000) \times 10}{100 \times 3}\right) = 1 - \Phi\left(\frac{10}{3}\right). \end{aligned}$$