

ΗΥ-217: Πιθανότητες - Χειμερινό Εξάμηνο 2008
Διδάσκων: Π. Τσακαλίδης

Λύσεις Πέμπτης Σειράς Ασκήσεων

Ημερομηνία Ανάθεσης: 31/10/2008

Ημερομηνία Παράδοσης: 12/11/2008

Άσκηση 1.

(α) Η συνάρτηση πιθανότητας της τ.μ. X είναι:

$$p_X(x) = \begin{cases} 1/9, & x = 2, \\ 2/9, & x = 3, \\ 3/9, & x = 4, \\ 2/9, & x = 5, \\ 1/9, & x = 6, \\ 0, & \text{αλλιώς} \end{cases}$$

η μέση τιμή:

$$E[X] = \sum_{x=1}^6 xp_X(x) = 2(1/9) + 3(2/9) + 4(3/9) + 5(2/9) + 6(1/9) = 4$$

και η διασπορά:

$$\text{var}[X] = \sum_{x=1}^6 (x - E[X])^2 p_X(x) = 4/3$$

(β) Έστω ότι $Z = X^2$, τότε η συνάρτηση πιθανότητας της τ.μ. Z είναι:

$$p_Z(z) = \begin{cases} 1/9, & z = 4, \\ 2/9, & z = 9, \\ 3/9, & z = 16, \\ 2/9, & z = 25, \\ 1/9, & z = 36, \\ 0, & \text{αλλιώς} \end{cases}$$

Η γραφική παράσταση της συνάρτησης πιθανότητας της τ.μ. Z φαίνεται στο σχήμα 1.

Άσκηση 2.

(α) Για να επιλέξουμε στην i -οστή προσπάθεια για πρώτη φορά την μπλέ μπάλα, τότε πρέπει να επιλέγουμε μια κόκκινη μπάλα στις $i - 1$ προσπάθειες και μετά να επιλέξουμε την μπλέ μπάλα.

$$P(X = 1) = 1/2, \quad P(X = 2) = (1/2)(1/3), \quad P(X = 3) = (1/2)(2/3)(1/4) = (1/3)(1/4), \dots$$

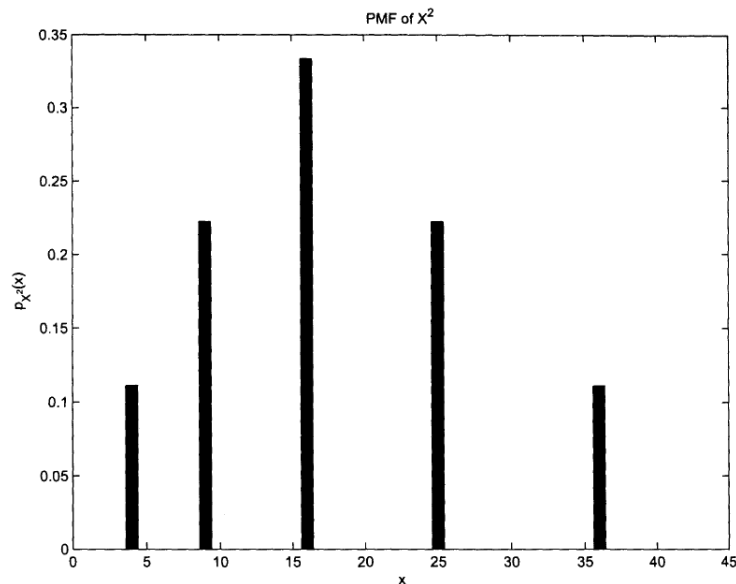
Γενικά, μπορούμε να αναγνωρίσουμε ότι:

$$P(X = i) = \frac{1}{i(i+1)}, \quad i \leq 1$$

(β) Για να υπολογίσουμε την πιθανότητα να μην επιλεγεί η μπλέ μπάλα κατά τις πρώτες n προσπάθειες, πρέπει να παρούμε το όριο καθώς το n τείνει στο άπειρο:

$$P(X > n) = (1/2)(2/3)(3/4)\dots(n/(n+1)) = 1/(n+1) \rightarrow 0 \text{ καθώς } n \rightarrow \infty$$

Συνεπώς η μπλέ μπάλα θα επιλεγεί τελικώς με πιθανότητα 1.



Σχήμα 1: Η γραφική παράσταση της συνάρτησης πιθανότητας της τ.μ. Z για την άσκηση 1.

(γ) Έχουμε ότι:

$$E[X] = \sum_i i P(X = i) = \sum_{i=1}^{\infty} i \left(\frac{1}{i(1+i)} \right) = \infty$$

Άσκηση 3.

(α) Γνωρίζουμε ότι το άθροισμα των πιθανοτήτων των δυνατών τιμών των x και y είναι 1. Συνεπώς:

$$1 \cdot 1 \cdot c + 1 \cdot 3 \cdot c + 2 \cdot 1 \cdot c + 2 \cdot 3 \cdot c + 4 \cdot 1 \cdot c + 4 \cdot 3 \cdot c = 1 \rightarrow c = \frac{1}{28}$$

(β) Υπάρχουν τρία σημεία για τα οποία ισχύει $Y < X$, τα $(1, 2)$, $(4, 1)$ και $(4, 3)$:

$$P(Y < X) = P((1, 2)) + P((4, 1)) + P((4, 3)) = 2 \cdot 1/28 + 4 \cdot 1/28 + 4 \cdot 3/28 = 18/28$$

(γ) Υπάρχουν δύο σημεία για τα οποία ισχύει $Y > X$, τα $(1, 3)$ και $(2, 3)$:

$$P(Y > X) = P((1, 3)) + P((2, 3)) = 1 \cdot 3/28 + 2 \cdot 3/28 = 9/28$$

(δ) Υπάρχει μόνο ένα σημείο για το οποίο ισχύει $Y = X$:

$$P(Y = X) = P((1, 1)) = 1 \cdot 1/28 = 1/28$$

Παρατηρήστε ότι το άθροισμα των πιθανοτήτων στο ερωτήματα (β), (γ) και (δ) είναι 1, όπως αναμένεται.

(ε) Υπάρχουν τρία σημεία για τα οποία ισχύει $y = 3$, τα $(1, 3)$, $(2, 3)$ και $(4, 3)$:

$$P(Y = 3) = P((1, 3)) + P((2, 3)) + P((4, 3)) = 3/28 + 6/28 + 12/28 = 21/28$$

(στ) Γενικά, για δύο διακριτές τυχαίες μεταβλητές X και Y για τις οποίες είναι ορισμένη η από κοινού συνάρτηση πιθανότητας, έχουμε:

$$p_X(x) = \sum_{y=-\infty}^{\infty} p_{X,Y}(x, y), \quad p_Y(y) = \sum_{x=-\infty}^{\infty} p_{X,Y}(x, y)$$

Στην δεδομένη περίπτωση το πλήθος των δυνατών ζευγών (X, Y) είναι μικρό, οπότε μπορούμε να υπολογίζουμε τις περιθωριακές συναρτήσεις πιθανότητας με απαρίθμηση. Για παράδειγμα:

$$p_X(2) = P((2, 1)) + P((2, 3)) = 8/28$$

Συνολικά έχουμε:

$$p_X(x) = \begin{cases} 4/28, & x = 1 \\ 8/28, & x = 2 \\ 16/28, & x = 4 \\ 0, & \text{αλλιώς} \end{cases}$$

και

$$p_Y(y) = \begin{cases} 7/28, & y = 1 \\ 21/28, & y = 3 \\ 0, & \text{αλλιώς} \end{cases}$$

(ζ) Γενικά, η μέση τιμή για οποιαδήποτε διακριτή τυχαία μεταβλητή Q είναι:

$$E[Q] = \sum_{x=-\infty}^{\infty} x p_Q(x)$$

. Για αυτό το πρόβλημα,

$$E[X] = 1(1/7) + 2(2/7) + 4(4/7) = 3$$

και

$$E[Y] = 1(1/4) + 3(3/4) = 5/2$$

(η) Η διασπορά για μια διακριτή τυχαία μεταβλητή Q υπολογίζεται ως $E[X^2] - E[X]^2$ ή ως $E[(X - E[X])^2]$. Χρησιμοποιώντας τον δεύτερο τρόπο έχουμε:

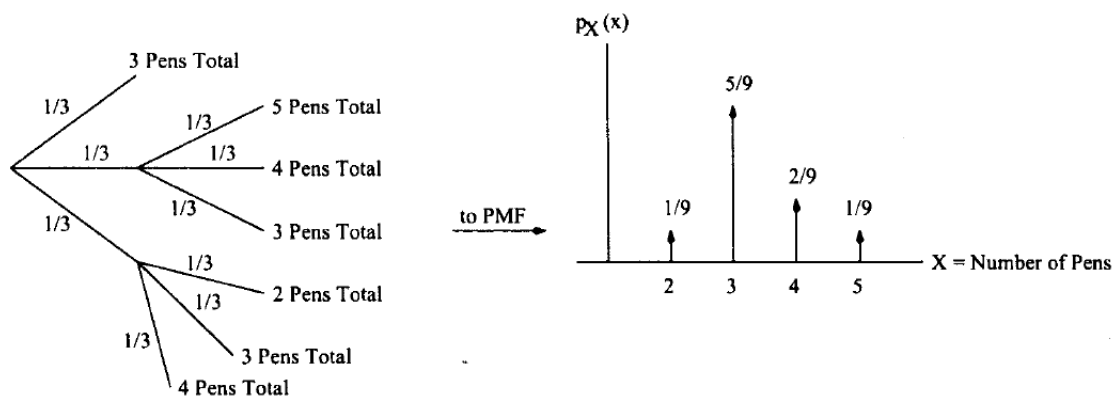
$$\text{var}(X) = (1 - 3)^2(1/7) + (2 - 3)^2(2/7) + (4 - 3)^2(4/7) = 10/7$$

$$\text{var}(Y) = (1 - 5/2)^2(1/4) + (3 - 5/2)^2(3/4) = 5/8$$

.

Άσκηση 4.

Η δένδρική αναπαράσταση και η συνάρτηση πυκνότητας πιθανότητας της τ.μ. X φαίνεται στο σχήμα 2.



Σχήμα 2: Η δένδρική αναπαράσταση και η συνάρτηση πυκνότητας πιθανότητας της τ.μ. X για την άσκηση 4.

(α) Από την η συνάρτηση πυκνότητας πιθανότητας της τ.μ. X έχουμε ότι: $P(A) = 5/9$.

(β)

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{2/9}{5/9} = \frac{2}{5}$$

(γ) Παρατηρήστε ότι η τ.μ. N είναι η ίδια με την τ.μ. X όπου ορίσαμε στο ερώτημα (α). Εφόσον έχουμε ήδη την συνάρτηση πυκνότητας πιθανότητας της N , μπορούμε να υπολογίσουμε την $E[N]$ άμεσα:

$$E[N] = \sum_n n p_N(n) = 2(1/9) + 3(5/9) + 4(2/9) + 5(1/9) = 10/3.$$

Άρα η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας $p_{N|C}(n)$ είναι:

$$p_{N|C}(n) = \begin{cases} 2/3, & n = 4 \\ 1/3, & n = 5 \\ 0, & \text{αλλιώς} \end{cases}$$

και

$$E[N|C] = \sum_n n p_{N|C}(n) = 4(2/3) + 5(1/3) = 13/3$$

(δ) Χρησιμοποιώντας της δεσμευμένη συνάρτηση πυκνότητας πιθανότητας $p_{N|C}(n)$, μπορούμε να υπολογίσουμε:

$$E[N^2|C] = \sum_n n^2 p_{N|C}(n) = 4(2/3) + 5^2(1/3) = 57/3.$$

Τότε: $\sigma_{N|C}^2 = E[N^2|C] - (E[N|C])^2 = 2/9$, άρα $\sigma_{N|C} = \sqrt{2}/3$.

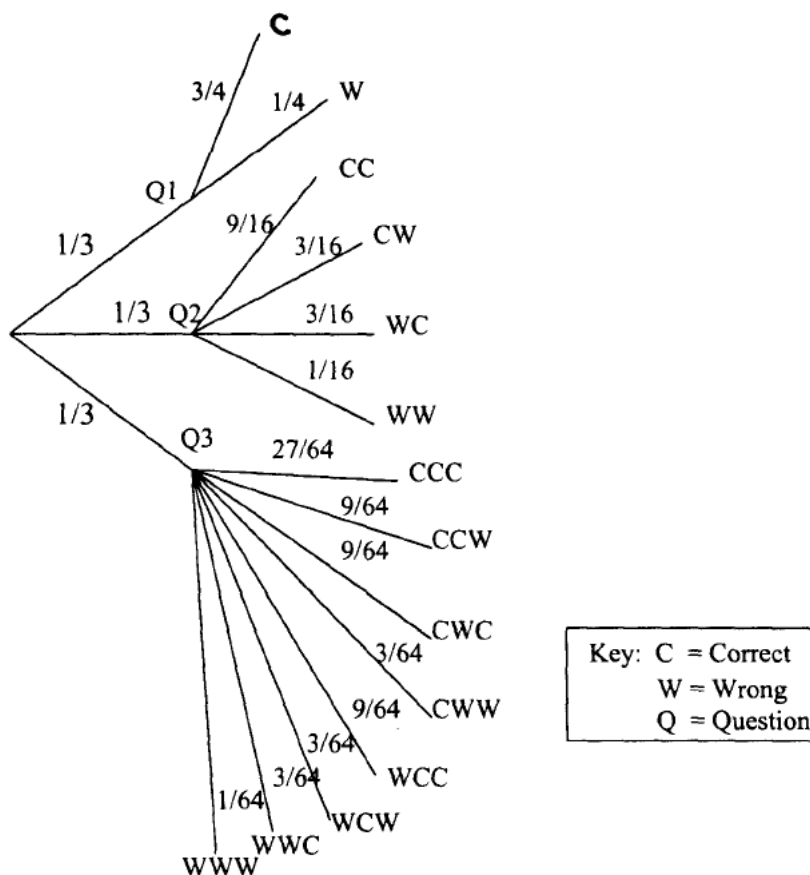
(ε) Η πιθανότητα ο Νίκος να πάρει πάνω από τρία μολύβια μέσα σε μια μέρα είναι:

$$p_N(4) + p_N(5) = 1/3.$$

Εφόσον οι κινήσεις του σε διαφορετικές μέρες είναι ανεξάρτητες: $P(D) = (1/3)^{16}$.

Άσκηση 5.

Το δενδρόγραμμα της άσκησης φαίνεται στο σχήμα 3.



Σχήμα 3: Το δενδρόγραμμα της άσκησης 5.

(α) Ο δειγματοχώρος της άσκησης χωρίζεται σε 3 περιοχές ανάλογα με το η May B.Right ερωτήθηκε

μία, δύο ή τρεις ερωτήσεις. Χρησιμοποιώντας το θεώρημα ολικής πιθανότητας έχουμε:

$$\begin{aligned} P(\text{May απαντάει λάθος σε όλες τις ερωτήσεις}) &= P(W|Q1)P(Q1) + P(WW|Q2)P(Q2) \\ &\quad + P(WWW|Q3)P(Q3) \\ &= 1/4 \times 1/3 + 1/16 \times 1/3 + 1/64 \times 1/3 \\ &= 21/192 \end{aligned}$$

(β) Χρησιμοποιώντας τον κανόνα δεσμευμένης πιθανότητας έχουμε:

$$\begin{aligned} P(Q3|\text{May απαντάει λάθος σε όλες τις ερωτήσεις}) &= \frac{P(Q3 \cap \text{May απαντάει λάθος σε όλες τις ερωτήσεις})}{P(\text{May απαντάει λάθος σε όλες τις ερωτήσεις})} \\ &= \frac{1/3 \times 1/64}{21/192} \\ &= \frac{1}{21} \end{aligned}$$

(γ) Η συνάρτηση πυκνότητας πιθανότητας της X είναι:

$$p_X(x) = \begin{cases} 1/3, & \text{αν } x = 1, 2, 3 \\ 0, & \text{αλλιώς} \end{cases}$$

η μέση τιμή της:

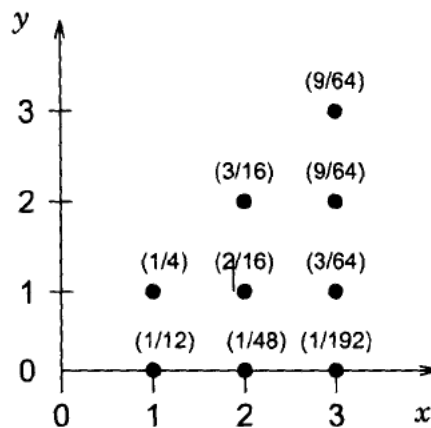
$$E[X] = 1(1/3) + 2(1/3) + 3(1/3) = 2$$

και για την διασπορά έχουμε:

$$E[X^2] = 1(1/3) + 4(1/3) + 9(1/3) = 14/3$$

$$\text{var}(X) = 14/3 - 4 = 2/3$$

(δ) Η από κοινού συνάρτηση πιθανότητας $p_{X,Y}(x, y)$ για τις τ.μ. X και Y φαίνεται στο σχήμα 4. Οι τιμές της $p_{X,Y}(x, y)$ υπολογίστηκαν χρησιμοποιώντας την σχέση: $p_{X,Y}(x, y) = p_{Y|X}(y|x)p_X(x)$.



Σχήμα 4: Η από κοινού συνάρτηση πιθανότητας $p_{X,Y}(x, y)$ για το ερώτημα (δ) της άσκησης 5.

(ε) Το επιμίσθιο μπορεί να οριστεί ως μια τυχαία μεταβλητή $B = 10X + 20Y$. Η μέση τιμή της B μπορεί να υπολογιστεί, αφού είναι η μέση τιμή ενός γραμμικού συνδιασμού δυο τ.μ. είναι ο γραμμικός συνδιασμός των μέσων τιμών των δύο τ.μ. Πρώτα υπολογίζουμε την μέση τιμή της Y . Από το ερώτημα (δ), μπορούμε να υπολογίσουμε την σ.π.π της Y χρησιμοποιώντας τον τύπο: $p_Y(y) = \sum_x p_{X,Y}(x, y)$, που μας δίνει:

$$p_Y(y) = \begin{cases} 7/64, & y = 0 \\ 27/64, & y = 1 \\ 21/64, & y = 2 \\ 9/64, & y = 3 \\ 0, & \text{αλλιώς} \end{cases}$$

Συνεπώς:

$$E[Y] = 0(7/64) + 1(27/64) + 2(21/64) + 3(9/64) = 96/64 = 3/2.$$

Οπότε:

$$E[B] = E[10X + 20Y] = 10E[X] + 20E[Y] = 50.$$

Για να υπολογίσουμε την διασπορά της B , χρησιμοποιούμε τον τύπο $var(B) = E[B^2] - (E[B])^2$. $E[B^2]$ μπορεί να υπολογιστεί αν παρατηρήσουμε ότι $B = g(X, Y)$, άρα:

$$\begin{aligned} E[B^2] &= \sum_x \sum_y (g(X, Y))^2 p_{X,Y}(x, y) = \sum_x \sum_y (10X + 20Y)^2 p_{X,Y}(x, y) \\ &= 100(1/12) + 900(1/4) + 400(1/48) + 1600(2/16) + 3600(3/16) + 900(1/192) + 2500(3/64) + 4900(9/64) + 8100(9/64) \\ var(B) &= E[B^2] - (E[B])^2 = 108800/192 \end{aligned}$$

(στ) Η τ.μ. Z ορίζεται ως το συνολικό πλήθος λάθος απαντήσεων της MBR σε μια σειρά 20 διαλέξεων. Οι διαλέξεις μεταξύ τους είναι ανεξάρτητες. Αυτό σημαίνει ότι το πλήθος των ερωτήσεων που απαντάει λάθος σε κάθε διάλεξη είναι ανεξάρτητο από το πλήθος των ερωτήσεων που απάντησε σε κάποια άλλη διάλεξη. Ομοίως, το πλήθος των ερωτήσεων που απαντά σε μια διάλεξη είναι ανεξάρτητο από το πλήθος των ερωτήσεων σε μια άλλη. Ας ορίσουμε την τ.μ. Q_n ως το πλήθος των ερωτήσεων που απαντά λάθος στην n -οστή διάλεξη. Συνεπώς $Z = Q_1 + Q_2 + \dots + Q_{20}$ και

$$E[Z] = E[Q_1] + E[Q_2] + \dots + E[Q_{20}]$$

επειδή όμως Q_n είναι ομοιόμορφα κατανομημένη σε κάθε διάλεξη, έχουμε:

$$E[Z] = 20E[Q_n].$$

Επειδή Q_n είναι το πλήθος των ερωτήσεων που απαντά λάθος στην n -οστή διάλεξη, $Q_n = X - Y = f(X, Y)$. Οπότε:

$$E[Q_n] = E[X - Y] = E[X] - E[Y] = 2 - 3/2 = 1/2$$

και

$$E[X] = 20E[Q_n] = 10.$$

Εφόσον οι τ.μ. Q_n είναι ανεξάρτητες και ομοιόμορφα κατανομημένες (i.i.d), έχουμε ότι:

$$var(Z) = var(Q_1 + Q_2 + \dots + Q_{20}) = 20var(Q_n).$$

Χρησιμοποιώντας την σχέση $var(Q_n) = E[Q_n^2] - (E[Q_n])^2$, έχουμε $E[Q_n^2]$. Όπως και στο ερώτημα (ε), όταν $B = g(X, Y)$, $Q_n = f(X, Y)$, τότε:

$$E[Q_n^2] = \sum_x \sum_y (f(X, Y))^2 p_{X,Y}(x, y) = \sum_x \sum_y (X - Y)^2 p_{X,Y}(x, y)$$

$$E[Q_n^2] = 1/12 + 2/16 + 9/64 + 4(1/48) + 4(3/64) + 9(1/192) = 2/3$$

$$var(Q_n) = 2/3 - 1/4 = 5/12$$

$$var(Z) = 20(5/12) = 100/12$$

Εναλλακτικά, μπορούμε να υπολογίσουμε την σ.π.π για την Q_n και να υπολογίσουμε μετά:

$$E[Q_n^2] = \sum_q q^2 p_{Q_n}(q)$$

Η σ.π.π για την Q_n μπορεί να υπολογιστεί από τον τύπο:

$$p_{Q_n}(q) = \sum_{x,y|f(x,y)=q} p_{X,Y}(x, y)$$

και είναι:

$$p_{Q_n}(q) = \begin{cases} 101/192, & q = 0 \\ 67/192, & q = 1 \\ 13/192, & q = 2 \\ 1/192, & q = 3 \\ 0, & \text{στηθερωσε} \end{cases}$$

Άσκηση 6.

(α) Χρησιμοποιούμε το γεγονός ότι $p_{N,K}(n, k) = p_{N|K}(n|k)p_K(k)$ για να υπολογίσουμε την από κοινού σ.π.π.:

$$p_{N,K}(n, k) = \begin{cases} 1/(4k), & k = 1, 2, 3, 4 \text{ ανδ } n = 1, \dots, k; \\ 0, & \text{αλλιώς} \end{cases}$$

(β) Η σ.π.π. $p_N(n)$ δίνεται από τον ακόλουθο τύπο:

$$p_N(n) = \sum_k p_{N,K}(n, k) = \sum_{k=n}^4 \frac{1}{4k}$$

δηλαδή:

$$p_N(n) = \begin{cases} 25/48, & n = 1 \\ 13/48, & n = 2 \\ 7/48, & n = 3 \\ 3/48, & n = 4 \\ 0, & \text{στηθερωσε} \end{cases}$$

(γ) Η δεσμευμένη σ.π.π. είναι:

$$p_{K|N}(k|2) = \frac{p_{N,K}(2, k)}{p_N(2)} = \begin{cases} 6/13, & k = 2; \\ 4/13, & k = 3; \\ 3/13, & k = 4; \\ 0, & \text{αλλιώς} \end{cases}$$

(δ) Έστω A το γεγονός ότι $2 \leq N \leq 3$. Πρώτα υπολογίζουμε την δεσμευμένη σ.π.π. του K δεδομένου του A :

$$p_{K|A}(k) = \frac{P(K = k, A)}{P(A)}$$

$$P(A) = p_N(2) + p_N(3) = 5/12$$

$$P(K = k, A) = \begin{cases} 1/8, & k = 2; \\ 1/12 + 1/12, & k = 3; \\ 1/16 + 1/16, & k = 4; \\ 0, & \text{αλλιώς} \end{cases}$$

$$p_{K|A}(k) = \begin{cases} 3/10, & k = 2; \\ 2/5, & k = 3; \\ 3/10, & k = 4; \\ 0, & \text{αλλιώς} \end{cases}$$

Επειδή η δεσμευμένη σ.π.π. του K δεδομένου του A είναι συμμετρική γύρω από το $k = 3$, γνωρίζουμε ότι $E[K|A] = 3$. Μπορούμε τώρα να υπολογίσουμε την δεσμευμένη διασπορά του K δεδομένου του A :

$$\text{var}(K|A) = E[(K - E[K|A])^2|A] = 3/10(2 - 3)^2 + 2/5(0) + 3/10(4 - 3)^2 = 3/5$$

(ε) Έστω C_i το κόστος ενός βιβλίου i και $E[C_i] = 3$. Έστω T το συνολικό κόστος, δηλαδή $T = C_1 + \dots + C_N$. Μπορούμε να υπολογίσουμε το $E[T]$ ως εξής:

$$E[T] = \sum_{i=1}^4 E[T|N = i]p_N(i) = E[C_i]p_N(1) + 2E[C_i]p_N(2) + 3E[C_i]p_N(3) + 4E[C_i]p_N(4)$$

$$= 3(25/48) + 6(13/48) + 9(7/48) + 12(1/16) = 21/4$$