

HY-217: Πιθανότητες -Χειμερινό Εξάμηνο 2012
Διδάσκων: Π. Τσακαλίδης

Λύσεις: Τέταρτη Σειρά Ασκήσεων

Ημερομηνία Ανάθεσης: 5/11/2012

Ημερομηνία Παράδοσης: 14/11/2012

Θέματα: Διακριτές Τυχαίες Μεταβλητές (II).

Άσκηση 1.

(α) Τα ζάρια είναι δίκαια, οπότε κάθε αριθμός έχει την ίδια πιθανότητα να εμφανιστεί: $\frac{1}{6}$. Συνεπώς,
 $p_X(8) = P(X=8) = P(\{1\}) = \frac{1}{6}$.
 $p_X(5) = P(X=5) = P(\{2,3\}) = P(\{2\}) + P(\{3\}) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}$.
 $p_X(-z) = P(X=-z) = P(\{4,5,6\}) = P(\{4\}) + P(\{5\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6}$.

(β) $-1 = \mu_X = \sum_x xp_X(x) = 8(\frac{1}{6}) + 5(\frac{2}{6}) + (-z)(\frac{3}{6}) = \frac{18-3z}{6}$. Λύνοντας ως προς z , έχουμε
 $z = \frac{-6-18}{-3} = 8$.

(γ) $E[\ln(|X|)] = \sum_x (\ln|x|)p_X(x) = (\ln 8)(\frac{1}{6}) + (\ln 5)(\frac{2}{6}) + (\ln z)(\frac{3}{6}) = \frac{\ln(200z^3)}{6} = \frac{\ln(102400)}{6} = 1.92$.

Άσκηση 2. Έστω B ο αριθμός των μπλέ σοκολάτων και Y ο αριθμός των κίτρινων τις οποίες παίρνει ο φίλος σας. Τότε το κόστος σε λεπτά είναι $X = 50B + 25Y$. Ο παρακάτω πίνακας δείχνει τα πιθανά (B, Y) ζεύγη, το πλήθος εμφάνισης κάθε ζεύγους $|(B, Y)|$, τη πιθανότητα του ζεύγους $P\{(B, Y)\} = |(B, Y)|/|\Omega|$ και τη τιμή που έχει η μεταβλητή X για το συγκεκριμένο ζεύγος. Προφανώς, υπάρχουν συνολικά $\binom{9}{3}$ τρόποι για να επιλέξει 3 σοκολάτες από τις 9, δηλαδή $|\Omega| = \binom{9}{3} = 84$.

(B, Y)	$ (B, Y) $	$P\{(B, Y)\}$	X
(0, 0)	$\binom{4}{3}$	$\frac{4}{84}$	0
(0, 1)	$\binom{3}{1}\binom{4}{2}$	$\frac{18}{84}$	25
(0, 2)	$\binom{3}{2}\binom{4}{1}$	$\frac{12}{84}$	50
(0, 3)	$\binom{3}{3}$	$\frac{1}{84}$	75
(1, 0)	$\binom{2}{1}\binom{4}{2}$	$\frac{12}{84}$	50
(1, 1)	$\binom{2}{1}\binom{3}{1}\binom{4}{1}$	$\frac{24}{84}$	75
(1, 2)	$\binom{2}{1}\binom{3}{2}$	$\frac{6}{84}$	100
(2, 0)	$\binom{2}{2}\binom{4}{1}$	$\frac{4}{84}$	100
(2, 1)	$\binom{2}{2}\binom{3}{1}$	$\frac{3}{84}$	125

Άρα, η X παίρνει τις τιμές 0, 25, 50, 75, 100, 125 με πιθανότητες:

$$\begin{aligned}p_X(0) &= P\{X=0\} = P\{(0,0)\} = \frac{4}{84}. \\p_X(25) &= P\{X=25\} = P\{(0,1)\} = \frac{18}{84}. \\p_X(50) &= P\{X=50\} = P\{(0,2) \cup (1,0)\} = P\{(0,2)\} + P\{(1,0)\} = \frac{12}{84} + \frac{12}{84} = \frac{24}{84}. \\p_X(75) &= P\{X=75\} = P\{(0,3) \cup (1,1)\} = P\{(0,3)\} + P\{(1,1)\} = \frac{1}{84} + \frac{24}{84} = \frac{25}{84}. \\p_X(100) &= P\{X=100\} = P\{(1,2) \cup (2,0)\} = P\{(1,2)\} + P\{(2,0)\} = \frac{6}{84} + \frac{4}{84} = \frac{10}{84}. \\p_X(125) &= P\{X=125\} = P\{(2,1)\} = \frac{3}{84}.\end{aligned}$$

$$\begin{aligned}E[2X+1] &= 2E[X] + 1 = 2\left(\frac{175}{3}\right) + 1 = \frac{353}{3} = 117.66. \\Var(2X+1) &= 2^2 Var(X) = 4\left(\frac{15625}{6}\right) = \frac{31250}{9} = 3472.22.\end{aligned}$$

Άσκηση 3.

(α) Έστω X ο αριθμός των χαμένων πακέτων σε ένα καρέ. Αναγνωρίζουμε ότι η τ.μ. X ακολουθεί διωνυμική κατανομή με παραμέτρους $n = 1000$ και $p = 0.001$, δηλαδή $X \sim \Delta(n = 1000, p = 0.001)$. Παρατηρούμε buffering όταν $X \geq 5$. Συνεπώς,

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - \sum_{k=0}^4 \binom{1000}{k} (0.001)^k (0.999)^{1000-k} = 0.003637$$

(β) Καθώς το n είναι μεγάλο και το p μικρό, μπορούμε να χρησιμοποιήσουμε την Poisson προσέγγιση της Διωνυμικής. Έστω Y μία Poisson τ.μ. με παραμέτρους $\lambda = np = 1000 \times 10^{-3} = 1$. Σύμφωνα με τη Poisson προσέγγιση, η πιθανότητα ότι πέντε ή περισσότερα πακέτα ενός καρέ θα χαθούν προσεγγίζεται από:

$$P(X \geq 5) \approx P(Y \geq 5) = 1 - P(Y \leq 4) = 1 - \sum_{k=0}^4 \frac{e^{-\lambda} \lambda^k}{k!} = 1 - e^{-1} \left(1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!}\right) = 0.003660.$$

Άσκηση 4.

(α,i) Προφανώς ο Κώστας λαμβάνει πίσω 100×100 ευρώ με πιθανότητα $1/50$ και 0 ευρώ με πιθανότητα $49/50$. Συνεπώς,

$$p_X(x) = \begin{cases} \frac{49}{50} & x = 0, \\ \frac{1}{50} & x = 10000, \\ 0 & \text{αλλού} \end{cases}$$

(α,ii) Η μέση τιμή είναι

$$E[X] = 0 \times \frac{49}{50} + 10000 \times \frac{1}{50} = 200.$$

Η δεύτερη ροπή είναι

$$E[X^2] = 0^2 \times \frac{49}{50} + 10000^2 \times \frac{1}{50} = 2 \times 10^6.$$

Συνεπώς, η X έχει διασπορά

$$\sigma_X^2 = E[X^2] - (E[X])^2 = 2 \times 10^6 - (200)^2 = 1960000.$$

και τυπική απόκλιση $\sigma_X = 1400$ ευρώ.

Στην ίδια απάντηση μπορούμε να φτάσουμε αν παρατηρήσουμε ότι $X = 10^4 Z$, όπου η Z ακολουθεί Bernoulli κατανομή με παράμετρο $p = 1/50$. Συνεπώς,

$$E[X] = 10^4 E[Z] = 10^4 p = 10^4 (1/50) = 200$$

και

$$\text{var}(X) = \text{var}(10^4 Z) = 10^8 \text{var}(Z) = 10^8 p(1-p) = 10^8 (1/50)(49/50) = 1960000$$

που δίνει μία τυπική απόκλιση ίση με $\sigma_X = \sqrt{\text{var}(X)} = 1400$ ευρώ.

(β, i) Σε αυτή την περίπτωση ο Κώστας λαμβάνει πίσω $Y = 100W$ ευρώ. Η τ.μ. W εκφράζει το πλήθος των λαχείων που κερδίζουν (από 100 ευρώ το καθένα) και ακολουθεί Διωνυμική κατανομή με παραμέτρους $n = 100$ και $p = 1/50$, $W \sim \Delta(n = 100, p = 1/50)$. Συνεπώς, η τ.μ. Y παίρνει τιμές στο σύνολο $\{0, 100, 200, \dots, 9800, 9900, 10000\}$ με πιθανότητα

$$p_Y(100k) = \binom{100}{k} \left(\frac{1}{50}\right)^k \left(\frac{49}{50}\right)^{100-k}; \quad k = 0, 1, 2, \dots, 100.$$

(β, ii) Έχουμε ότι η με τη νέα στρατηγική, η μέση τιμή είναι

$$E[Y] = 100E[W] = 100np = 100 \times 100 \times \frac{1}{50} = 200,$$

η διασπορά είναι

$$\text{var}(Y) = \text{var}(100W) = 10^4 \text{var}(W) = 10^4 np(1-p) = 10^4 100(1/50)(49/50) = 196000$$

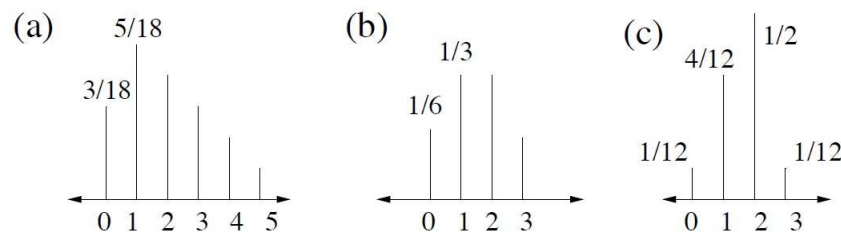
που δίνει τυπική απόκλιση ίση με $\sigma_Y = 140$ ευρώ. Και οι δύο στρατηγικές δίνουν το ίδιο μέσο κέρδος αλλά η δεύτερη στρατηγική έχει 10 φορές μικρότερη τυπική απόκλιση.

Άσκηση 5. [Απόσταση μεταξύ δύο τυχαία επιλεγμένων κορυφών]

(α) Υπάρχουν 36 πιθανές τιμές των (i, j) και $D = |i - j|$. Αρχικά,

$$p_D(0) = P\{D = 0\} = \frac{|\{(1,1), \dots, (6,6)\}|}{36} = \frac{6}{36} = \frac{3}{18} = \frac{1}{6}.$$

Για $1 \leq k \leq 5$, $p_D(k) = P\{D = k\} = \frac{|\{(1,k+1), \dots, (6-k,6)\} \cup \{(k+1,1), \dots, (6,6-k)\}|}{36} = \frac{2(6-k)}{36} = \frac{6-k}{18}$ και $p_D(k) = 0$ για $k \notin \{0, \dots, 5\}$.



$$\begin{aligned} E[D] &= \frac{3}{18} \cdot 0 + \frac{5}{18} \cdot 1 + \frac{4}{18} \cdot 2 + \frac{3}{18} \cdot 3 + \frac{2}{18} \cdot 4 + \frac{1}{18} \cdot 5 = \frac{35}{18} = 1.944. \\ E[D^2] &= \frac{3}{18} \cdot 0^2 + \frac{5}{18} \cdot 1^2 + \frac{4}{18} \cdot 2^2 + \frac{3}{18} \cdot 3^2 + \frac{2}{18} \cdot 4^2 + \frac{1}{18} \cdot 5^2 = \frac{105}{18} = 5.8333. \\ \text{var}(D) &= E[D^2] - (E[D])^2 = 2.0525. \end{aligned}$$

(β) Έστω $i = 1$ χωρίς βλάβη της γενικότητας. Τότε $D = 0$ αν $j = 1$, $D = 1$ αν $j \in \{2, 6\}$, $D = 2$ αν $j \in \{3, 5\}$, και $D = 3$ αν $j = 4$. Επομένως, $p_D(0) = p_D(3) = \frac{1}{6}$, $p_D(1) = p_D(2) = \frac{2}{6} = \frac{1}{3}$, και $p_D(k) = 0$ αν $k \notin \{0, 1, 2, 3\}$.

$$\begin{aligned} E[D] &= \frac{1}{6} \cdot 0 + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 2 + \frac{1}{6} \cdot 3 = 1.5. \\ E[D^2] &= \frac{1}{6} \cdot 0 + \frac{1}{3} \cdot 1^2 + \frac{1}{3} \cdot 2^2 + \frac{1}{6} \cdot 3^2 = \frac{19}{6}. \\ \text{var}(D) &= E[D^2] - (E[D])^2 = \frac{11}{12} = 0.9167. \end{aligned}$$

(γ) Έστω $i = 1$ χωρίς βλάβη της γενικότητας. Τότε $D = 0$ αν $j = 1$, $D = 1$ αν $j \in \{2, 5, 9, 12\}$, και $D = 3$ αν $j = 7$. Για τις άλλες 6 επιλογές του j , $D = 2$. Συνεπώς, $p_D(0) = p_D(3) = \frac{1}{12}$, $p_D(1) = \frac{4}{12} = \frac{1}{3}$ και $p_D(2) = \frac{6}{12} = \frac{1}{2}$. Σε αυτή την περίπτωση βρίσκουμε ότι $E[D] = \frac{19}{12} = 1.5833$ και $\text{var}(D) = 0.5764$.

Άσκηση 6.

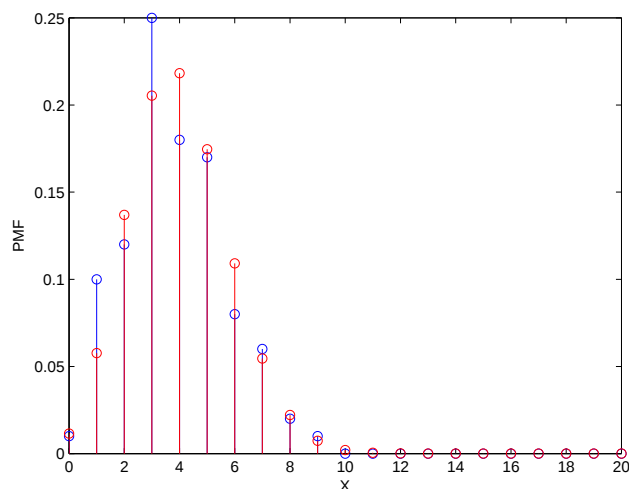
(α)

```
>> N = 20;
>> p = 0.2;
>> X = sum(rand(20, 1) < 0.2);
```

(β)

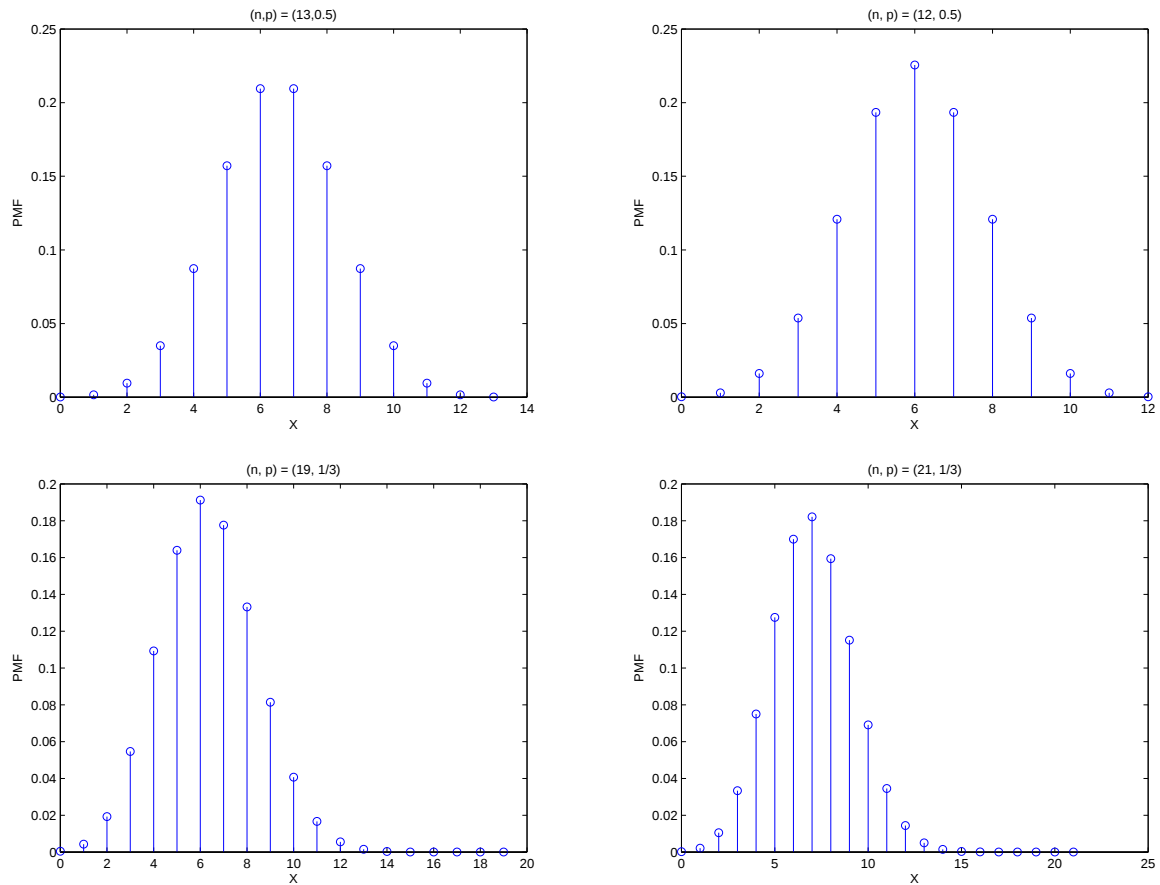
```
>> allX = zeros(100,1);
>> for i=1:100
>>   allX(i) = sum(rand(20,1) < 0.2);
>> end
>> estPMF = zeros(21,1);
>> for value=0:20
>>   estPMF(value+1)=sum(allX == value); %the +1 to avoid index 0!
>> end
>> estPMF = estPMF/100; % normalization
>> allvalues = 0:20; % 21 x 1 vector;
>> stem(allvalues,estPMF);
>> hold on;
>> for value=0:20
>>   PMF(value+1)=(factorial(N)/(factorial(value)*
>>       factorial(N-value)))*(p^value)*(1-p)^(N-value);
>> end
>> stem(allvalues, PMF, 'r')
```

Στο σχήμα 1 φαίνεται η εκτιμώμενη (μπλε) και η θεωρητική (κόκκινο) $\sigma.π.$ της τυχαίας μεταβλητής X . Η X είναι διωνυμική τυχαία μεταβλητή με παραμέτρους $n = 20$ και $p = 0.2$.



Σχήμα 1: Η $\sigma.π.$ της τυχαίας μεταβλητής X για το ερώτημα (β) της Άσκησης 6.

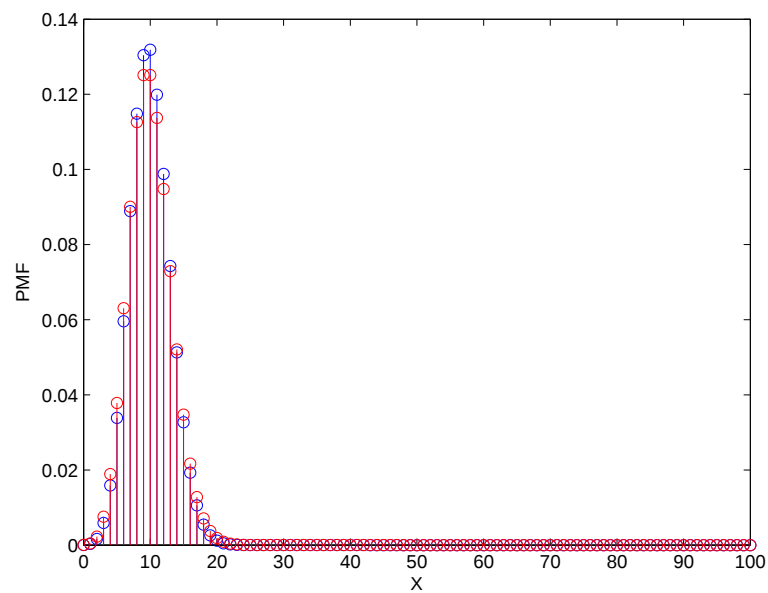
(γ) Για $p = 1/2$, η $\sigma.π.$ είναι συμμετρική. Τότε όταν np είναι ακέραιος, η συνάρτηση μεγιστοποιείται στο σημείο np . Όταν $p \neq 1/2$, η παράσταση δεν είναι συμμετρική. Τότε όταν np είναι ακέραιος, η συνάρτηση μεγιστοποιείται στο np , ενώ όταν το np δεν είναι ακέραιος τότε οι δύο μέγιστες τιμές της συνάρτησης βρίσκονται αριστερά και δεξιά του np . Για τις γραφικές παραστάσεις βλέπε το Σχήμα 2.



Σχήμα 2: Άσκηση 6 ερώτημα (γ).

(δ)

```
>> n=100; p=0.1;
>> for value=0:n
>>   BinomialPMF(value+1)=(factorial(n)/(factorial(value)*factorial(n-value)))
>>   *(p^value)*((1-p)^(n-value));
>> end
>> allvalues = 0:n;
>> clf;
>> stem(allvalues ,BinomialPMF);
>> hold on;
>> for value=0:n
>>   PoissonPMF(value+1)=( (10)^value)*exp(-10)/factorial(value);
>> end
>> stem(allvalues,PoissonPMF,'r'); % pretty good match!
```



Σχήμα 3: Άσκηση 6 ερώτημα (δ). Με μπλε είναι η γραφική παράσταση της διωνυμικής για $(n = 100, p = 0.1)$ και με κόκκινο η γραφική παράσταση της Poisson με παράμετρο $\lambda = 10$.