Κεφάλαιο 2

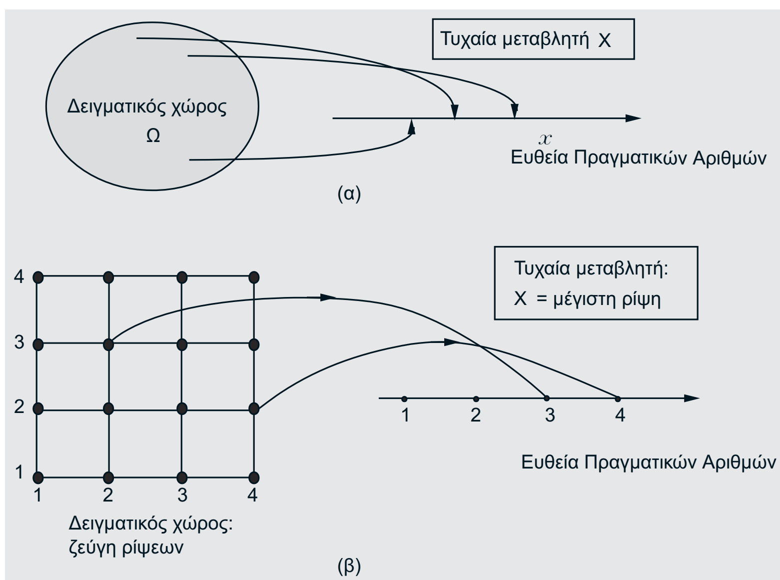
Διακριτές Τυχαίες Μεταβλητές

Περιεχόμενα

2.1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Πολλά μοντέλα πιθανοτήτων έχουν αριθμητικά αποτελέσματα, π.χ. όταν αντιστοιχούν σε μετρήσεις οργάνων, ή τιμές μετοχών στο χρηματιστήριο. Άλλα μοντέλα μπορεί να μην έχουν αριθμητικά αποτελέσματα, αλλά ενδέχεται να συνδέονται με αριθμητικές τιμές που μας ενδιαφέρουν. Για παράδειγμα, εάν ένα πείραμα αφορά την επιλογή φοιτητών από έναν πληθυσμό, ίσως να ενδιαφερόμαστε για το μέσο όρο της βαθμολογίας τους. Πιθανότητες που σχετίζονται με τέτοιου είδους αριθμητικές τιμές μπορούν να υπολογισθούν χρησιμοποιώντας την έννοια της **τυχαίας μεταβλητής** που αποτελεί το αντικείμενο του κεφαλαίου αυτού.

Δεδομένου ενός πειράματος και του αντίστοιχου συνόλου των δυνατών αποτελεσμάτων (το δειγματικό χώρο), μία τυχαία μεταβλητή συνδέει ένα συγκεκριμένο αριθμό με κάθε αποτέλεσμα, βλ. Σχ. 2.1. Αναφερόμαστε στον αριθμό αυτό ως την **αριθμητική τιμή** ή απλώς την **τιμή** της τυχαίας μεταβλητής. Από μαθηματικής άποψης, **η τυχαία μεταβλητή είναι μία συνάρτηση πραγματικών τιμών του πειραματικού αποτελέσματος**.



Σχήμα 2.1: (α) Απεικόνιση μίας τυχαίας μεταβλητής. Είναι μία συνάρτηση, η οποία συνδέει μία αριθμητική τιμή με κάθε δυνατό αποτέλεσμα του πειράματος. (β) Παράδειγμα μίας τυχαίας μεταβλητής. Το πείραμα αποτελείται από 2 ρίψεις ενός τετράεδρου ζαριού και η τυχαία μεταβλητή είναι το μέγιστο των δύο ρίψεων. Εάν το αποτέλεσμα του πειράματος είναι (4,2), η τιμή της τυχαίας μεταβλητής είναι 4.

Παρουσιάζουμε μερικά παραδείγματα τυχαίων μεταβλητών:

(α) Σε ένα πείραμα το οποίο αποτελείται από 5 ρίψεις ενός νομίσματος, το πλήθος των κορωνών στην ακολουθία των ρίψεων είναι μία τυχαία μεταβλητή. Ωστόσο, η μήκους 5 ακολουθία των κορωνών και γραμμάτων δεν θεωρείται τυχαία μεταβλητή, διότι δεν έχει μία συγκεκριμένη αριθμητική τιμή.

(β) Σε ένα πείραμα το οποίο αποτελείται από δυο ρίψεις ενός ζαριού, τα παρακάτω συνιστούν παραδείγματα τυχαίων μεταβλητών:

  (i) Το άθροισμα των ρίψεων.

  (ii) Ο αριθμός των 6 στις δύο ρίψεις.

  (iii) Η δεύτερη ρίψη υψωμένη στη δύναμη του πέντε.

(γ) Σε ένα πείραμα το οποίο αποτελείται από την αποστολή ενός μηνύματος, ο χρόνος που χρειάζεται να αποσταλεί το μήνυμα, το πλήθος των συμβόλων που λαμβάνονται λάθος και η καθυστέρηση με την οποία το μήνυμα ελήφθη, είναι τυχαίες μεταβλητές.

Υπάρχουν αρκετές βασικές έννοιες οι οποίες συνδέονται με τυχαίες μεταβλητές και οι οποίες συνοψίζονται παρακάτω. Στο κεφάλαιο αυτό θα συζητήσουμε λεπτομερώς τις έννοιες αυτές.

Κύριες Έννοιες Τυχαίων Μεταβλητών

Ας αρχίσουμε με ένα μοντέλο πιθανότητας ενός πειράματος:

- Μία **τυχαία μεταβλητή** είναι μία συνάρτηση πραγματικών τιμών του αποτελέσματος του πειράματος.
- Μία **συνάρτηση μίας τυχαίας μεταβλητής** ορίζει μία νέα τυχαία μεταβλητή.
- Μπορούμε να συνδέσουμε με κάθε τυχαία μεταβλητή κάποιους "μέσους όρους" που μας ενδιαφέρουν, όπως τη **μέση τιμή** και τη **διασπορά**.
- Μία τυχαία μεταβλητή μπορεί να **δεσμευθεί** από ένα γεγονός, ή μία άλλη τυχαία μεταβλητή.
- Υπάρχει η έννοια της **ανεξαρτησίας** μίας τυχαίας μεταβλητής από ένα γεγονός ή από μία άλλη τυχαία μεταβλητή.

Μία τυχαία μεταβλητή λέγεται **διακριτή** εάν το πεδίο τιμών της (το σύνολο των τιμών που μπορεί να πάρει) είναι πεπερασμένο ή αριθμήσιμα άπειρο. Για παράδειγμα, οι τυχαίες μεταβλητές στις οποίες αναφερθήκαμε στα (α) και (β) παραπάνω μπορούν να πάρουν ένα πεπερασμένο πλήθος αριθμητικών τιμών και επομένως είναι διακριτές.

Μία τυχαία μεταβλητή που μπορεί να πάρει άπειρο και μη αριθμήσιμο πλήθος τιμών δεν είναι διακριτή. Για παράδειγμα, θεωρήστε το πείραμα επιλογής ενός σημείου a στο διάστημα $[-1, 1]$. Η τυχαία μεταβλητή η οποία συνδέει την αριθμητική τιμή a^2 με το αποτέλεσμα a δεν είναι διακριτή. Ωστόσο, η τυχαία μεταβλητή η οποία συνδέει το a με την αριθμητική τιμή

$$\text{sgn}(a) = \begin{cases} 1, & \text{εάν } a > 0, \\ 0, & \text{εάν } a = 0, \\ -1, & \text{εάν } a < 0, \end{cases}$$

είναι διακριτή.

Στο κεφάλαιο αυτό, εστιάζουμε αποκλειστικά σε διακριτές τυχαίες μεταβλητές, παρόλο που θα παραλείπουμε τον όρο "διακριτός".

Έννοιες Διακριτών Τυχαίων Μεταβλητών

Ας αρχίσουμε με ένα μοντέλο πιθανότητας ενός πειράματος.

- Μία **διακριτή τυχαία μεταβλητή** είναι μία πραγματική συνάρτηση του αποτελέσματος του πειράματος η οποία μπορεί να πάρει ένα πεπερασμένο ή αριθμήσιμα άπειρο πλήθος τιμών.
- Μία διακριτή τυχαία μεταβλητή συνδέεται με μία **συνάρτηση μάζας πιθανότητας (ΣΜΠ)**, η οποία ορίζει την πιθανότητα κάθε αριθμητικής τιμής που μπορεί να πάρει η τυχαία μεταβλητή.
- Μία **συνάρτηση μίας διακριτής τυχαίας μεταβλητής** ορίζει μία άλλη τυχαία μεταβλητή, της οποίας η ΣΜΠ καθορίζεται από τη ΣΜΠ της αρχικής τυχαίας μεταβλητής.

Θα συζητήσουμε κάθε μία από τις παραπάνω έννοιες και τη σχετική μεθοδολογία στις παραγράφους που ακολουθούν. Επιπλέον, θα δώσουμε παραδείγματα τυχαίων μεταβλητών οι οποίες απαντώνται συχνά. Στο Κεφάλαιο 3 θα συζητήσουμε γενικές (όχι απαραίτητα διακριτές) τυχαίες μεταβλητές.

Παρόλο που στο κεφάλαιο αυτό φαίνεται να καλύπτονται πολλές καινούργιες έννοιες, στη πραγματικότητα πρόκειται για κάτι απλούστερο. Η τακτική που θα ακολουθήσουμε είναι απλώς να πάρουμε τις έννοιες του 1ου κεφαλαίου (πιθα-

νότητες, δεσμευμένες πιθανότητες, ανεξαρτησία, κλπ.) και να τις εφαρμόσουμε για τυχαίες μεταβλητές αντί για γεγονότα, μαζί με μερικούς καινούργιους συμβολισμούς. Οι μόνες πραγματικά καινούργιες έννοιες αφορούν μέσες τιμές και διασπορές.

2.2 ΣΥΝΑΡΤΗΣΕΙΣ ΜΑΖΑΣ ΠΙΘΑΝΟΤΗΤΑΣ

Ο σημαντικότερος τρόπος να χαρακτηρίσουμε μία τυχαία μεταβλητή είναι να προσδιορίσουμε τις πιθανότητες των τιμών που μπορεί να πάρει. Για μία διακριτή τυχαία μεταβλητή X, οι τιμές αυτές εκφράζονται από τη **Συνάρτηση Μάζας Πιθανότητας** (ΣΜΠ για συντομία) της X, η οποία συμβολίζεται με p_X. Συγκεκριμένα, εάν x είναι οποιαδήποτε δυνατή τιμή της X, η **μάζα πιθανότητας** του x, η οποία συμβολίζεται με $p_X(x)$, είναι η πιθανότητα του γεγονότος $\{X = x\}$ το οποίο αποτελείται από όλα τα δυνατά αποτελέσματα τα οποία αντιστοιχούν σε τιμή της X ίση με το x:

$$p_X(x) = \mathrm{P}(\{X = x\}).$$

Για παράδειγμα, εάν ένα πείραμα αποτελείται από δυο ανεξάρτητες ρίψεις ενός αμερόληπτου νομίσματος και εάν X είναι ο αριθμός των κορωνών που προκύπτουν, τότε η ΣΜΠ της X είναι

$$p_X(x) = \begin{cases} 1/4, & \text{εάν } x = 0 \text{ ή } x = 2, \\ 1/2, & \text{εάν } x = 1, \\ 0, & \text{αλλιώς.} \end{cases}$$

Σε ό,τι ακολουθεί, θα παραλείπουμε τις αγκύλες από το συμβολισμό γεγονότων/συνόλων, εφόσον δεν υπάρχει ασάφεια. Συγκεκριμένα, θα γράφουμε συνήθως $\mathrm{P}(X = x)$ αντί του πιο σωστού $\mathrm{P}(\{X = x\})$. Επίσης, θα τηρήσουμε την εξής σύμβαση: **θα χρησιμοποιούμε κεφαλαία γράμματα για να δηλώσουμε τυχαίες μεταβλητές και μικρά γράμματα για να δηλώσουμε πραγματικούς αριθμούς, όπως αριθμητικές τιμές μίας τυχαίας μεταβλητής.**

Σημειώστε ότι

$$\sum_x p_X(x) = 1,$$

όπου στο παραπάνω άθροισμα οι τιμές x κυμαίνονται σε όλες τις δυνατές αριθμητικές τιμές της X. Το άθροισμα ισούται με 1 και αυτό συμπεραίνεται από τα αξιώματα της προσθετικότητας και κανονικοποίησης, επειδή τα γεγονότα $\{X = x\}$ είναι ξένα μεταξύ τους και δημιουργούν μία διαμέριση του δειγματικού χώρου, καθώς το x παίρνει όλες τις δυνατές τιμές της X. Παρόμοια, για οποιοδήποτε

υποσύνολο S δυνατών τιμών της X, έχουμε

$$\mathrm{P}(X \in S) = \sum_{x \in S} p_X(x).$$

Για παράδειγμα, εάν X είναι ο αριθμός των κορωνών που προκύπτουν σε δύο ανεξάρτητες ρίψεις ενός αμερόληπτου νομίσματος, όπως παραπάνω, η πιθανότητα τουλάχιστον μίας κορώνας είναι

$$\mathrm{P}(X > 0) = \sum_{x=1}^{2} p_X(x) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$$

Ο υπολογισμός της ΣΜΠ της X είναι εννοιολογικά απλός και απεικονίζεται στο παράδειγμα του Σχ. 2.2.

Υπολογισμός της ΣΜΠ μίας Τυχαίας Μεταβλητής X

Για κάθε δυνατή τιμή x της X:

1. Συλλέγουμε όλα τα δυνατά αποτελέσματα που αντιστοιχούν στο γεγονός $\{X = x\}$.

2. Προσθέτουμε τις πιθανότητές τους για να πάρουμε την $p_X(x)$.

Η Τυχαία Μεταβλητή Bernoulli

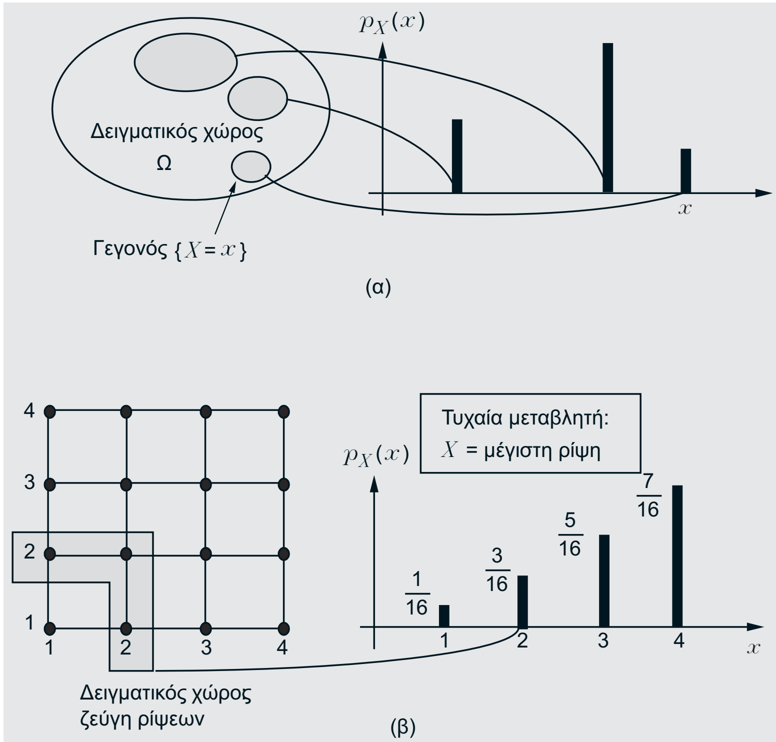
Θεωρήστε τη ρίψη ενός νομίσματος η οποία είναι κορώνα με πιθανότητα p και γράμματα με πιθανότητα $1 - p$. Η τυχαία μεταβλητή **Bernoulli** παίρνει τις δύο τιμές 1 και 0, ανάλογα με το εάν το αποτέλεσμα είναι κορώνα ή γράμματα:

$$X = \begin{cases} 1, & \text{εάν είναι κορώνα,} \\ 0, & \text{εάν είναι γράμματα.} \end{cases}$$

Η ΣΜΠ της είναι

$$p_X(x) = \begin{cases} p, & \text{εάν } x = 1, \\ 1 - p, & \text{εάν } x = 0. \end{cases}$$

Παρόλη την απλότητά της, η τυχαία μεταβλητή Bernoulli είναι πολύ σημαντική. Στην πράξη χρησιμοποιείται για να περιγράψει καταστάσεις όπου υπάρχουν μόνο δύο αποτελέσματα, όπως:



Σχήμα 2.2: (α) Απεικόνιση της μεθόδου υπολογισμού της ΣΜΠ μίας τυχαίας μεταβλητής X. Για κάθε δυνατή τιμή x, συλλέγουμε όλα τα αποτελέσματα τα οποία αντιστοιχούν στο $X = x$ και προσθέτουμε τις πιθανότητές τους για να πάρουμε την $p_X(x)$. (β) Υπολογισμός της ΣΜΠ p_X της τυχαίας μεταβλητής X= μέγιστο δύο ανεξάρτητων ρίψεων ενός 4-εδρου ζαριού. Υπάρχουν 4 δυνατές τιμές x: $1, 2, 3, 4$. Για να υπολογίσουμε την $p_X(x)$ για κάποιο συγκεκριμένο x, προσθέτουμε τις πιθανότητες των αποτελεσμάτων που αντιστοιχούν στο x. Για παράδειγμα, υπάρχουν 3 δυνατά αποτελέσματα τα οποία αντιστοιχούν στο $x = 2$: $(1, 2), (2, 2), (2, 1)$. Καθένα από αυτά τα δυνατά αποτελέσματα έχει πιθανότητα $1/16$, επομένως $p_X(2) = 3/16$, όπως φαίνεται στο σχήμα.

(α) Την κατάσταση μίας τηλεφωνικής γραμμής, η οποία μπορεί να είναι διαθέσιμη ή μη διαθέσιμη σε κάποια χρονική στιγμή.

(β) Την υγεία ενός ατόμου το οποίο μπορεί να είναι υγιές ή άρρωστο με κάποια ασθένεια.

(γ) Την προτίμηση ενός ατόμου, η οποία μπορεί να είναι η υποστήριξη ή όχι ενός πολιτικού υποψηφίου.

Επιπλέον, συνδυάζοντας πολλές τυχαίες μεταβλητές Bernoulli, μπορούμε να κατασκευάσουμε πιο πολύπλοκες τυχαίες μεταβλητές, όπως τη διωνυμική τυχαία μεταβλητή την οποία συζητούμε ακολούθως.

Η Διωνυμική Τυχαία Μεταβλητή

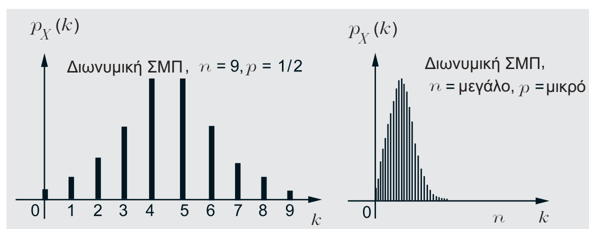
Ένα νόμισμα ρίχνεται n φορές. Σε κάθε ρίψη το νόμισμα είναι κορώνα με πιθανότητα p και γράμματα με πιθανότητα $1 - p$, ανεξαρτήτως των προηγούμενων ρίψεων. Έστω X το πλήθος των κορωνών σε n ρίψεις. Αναφερόμαστε στη X σαν μία **διωνυμική** τυχαία μεταβλητή με **παραμέτρους** n **και** p. Η ΣΜΠ της X αποτελείται από τις διωνυμικές πιθανότητες που υπολογίστηκαν στην Παράγραφο 1.5:

$$p_X(k) = \mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

(Σημειώνουμε ότι απλοποιούμε το συμβολισμό και χρησιμοποιούμε το k, αντί του x, για να δηλώσουμε ακέραιες τιμές τυχαίων μεταβλητών.) Η ιδιότητα της κανονικοποίησης, στην περίπτωση της διωνυμικής τυχαίας μεταβλητής, δίνει

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1.$$

Μερικές ειδικές περιπτώσεις της διωνυμικής ΣΜΠ απεικονίζονται στο Σχ. 2.3.



Σχήμα 2.3: Η ΣΜΠ μίας διωνυμικής τυχαίας μεταβλητής. Εάν $p = 1/2$, η ΣΜΠ είναι συμμετρική γύρω από το $n/2$. Διαφορετικά η ΣΜΠ κλίνει προς το μηδέν 0 εάν $p < 1/2$ και προς το n εάν $p > 1/2$.

Η Γεωμετρική Τυχαία Μεταβλητή

Υποθέστε ότι ρίχνουμε επανειλημμένα και ανεξάρτητα ένα νόμισμα με πιθανότητα κορώνας ίση με p, όπου $0 < p < 1$. Η **γεωμετρική** τυχαία μεταβλητή είναι το πλήθος X των ρίψεων μέχρι την εμφάνιση της πρώτης κορώνας. Η ΣΜΠ είναι

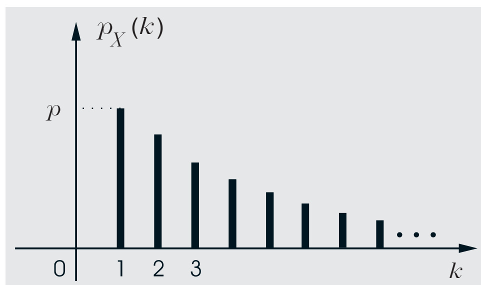
$$p_X(k) = (1-p)^{k-1} p, \qquad k = 1, 2, \ldots,$$

όπου $(1-p)^{k-1}p$ είναι η πιθανότητα της ακολουθίας που αποτελείται από $k-1$ διαδοχικές αποτυχίες (γράμματα) ακολουθούμενες από μία επιτυχία (κορώνα)· βλ.

Σχ. 2.4. Αυτή είναι μία αποδεκτή ΣΜΠ διότι

$$\sum_{k=1}^{\infty} p_X(k) = \sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{k=0}^{\infty}(1-p)^k = p \cdot \frac{1}{1-(1-p)} = 1.$$

Φυσικά, εδώ χρησιμοποιούμε ρίψεις νομισμάτων απλώς για να διευκολύνουμε την κατανόηση. Γενικότερα μπορούμε να ερμηνεύσουμε τη γεωμετρική τυχαία μεταβλητή σαν τον αριθμό από επαναλαμβανόμενες ανεξάρτητες δοκιμές, μέχρι την πρώτη "επιτυχία". Κάθε δοκιμή έχει πιθανότητα επιτυχίας p και το πλήθος των δοκιμών μέχρι (και συμπεριλαμβανομένης) της πρώτης επιτυχίας μοντελοποιείται από τη γεωμετρική τυχαία μεταβλητή. Η έννοια της "επιτυχίας" εξαρτάται από τη συγκεκριμένη εφαρμογή. Για παράδειγμα, μπορεί να εννοεί την επιτυχία σε ένα τεστ μετά από κάποιες αποτυχημένες προσπάθειες ή την εύρεση κάποιου αντικειμένου μετά από έρευνα ή την επιτυχή πρόσβαση σε έναν υπολογιστή σε μία δεδομένη προσπάθεια, κλπ.



Σχήμα 2.4: Η ΣΜΠ

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots,$$

μίας γεωμετρικής τυχαίας μεταβλητής. Είναι φθίνουσα γεωμετρική πρόοδος με παράμετρο $1-p$.

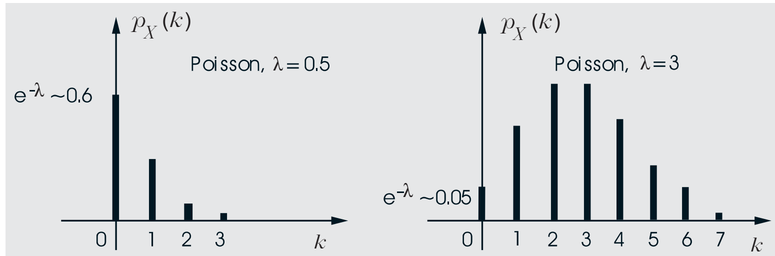
Η Τυχαία Μεταβλητή Poisson

Μία τυχαία μεταβλητή Poisson περιγράφεται από τη ΣΜΠ

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots,$$

όπου λ είναι μία θετική παράμετρος που χαρακτηρίζει την ΣΜΠ, βλ. Σχ. 2.5. Είναι μία αποδεκτή ΣΜΠ διότι

$$\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right) = e^{-\lambda}e^{\lambda} = 1.$$



Σχήμα 2.5: Η ΣΜΠ $e^{-\lambda}\lambda^k/k!$ μίας τυχαίας μεταβλητής Poisson για διαφορετικές τιμές του λ. Παρατηρήστε ότι εάν $\lambda < 1$ τότε η ΣΜΠ είναι φθίνουσα, ενώ εάν $\lambda > 1$ η ΣΜΠ αρχικά είναι αύξουσα και μετά είναι φθίνουσα, καθώς το k αυξάνει (αυτό αποδεικνύεται στα προβλήματα στο τέλος του κεφαλαίου).

Για να αντιληφθείτε καλύτερα την τυχαία μεταβλητή Poisson, σκεφθείτε μία διωνυμική τυχαία μεταβλητή με πολύ μικρό p και πολύ μεγάλο n. Για παράδειγμα, θεωρήστε τον αριθμό λαθών ενός βιβλίου που περιέχει n λέξεις, όπου η πιθανότητα p ότι οποιαδήποτε δεδομένη λέξη είναι ανορθόγραφη είναι πολύ μικρή (συνδέστε μία λέξη με τη ρίψη ενός νομίσματος το οποίο είναι κορώνα όταν η λέξη είναι ανορθόγραφη) ή τον αριθμό αυτοκινήτων τα οποία εμπλέκονται σε ατυχήματα σε μία πόλη κάποια μέρα (συνδέστε ένα αυτοκίνητο με τη ρίψη ενός νομίσματος το οποίο είναι κορώνα όταν το αυτοκίνητο έχει ατύχημα). Τέτοιες τυχαίες μεταβλητές μπορούν να μοντελοποιηθούν εύστοχα από την Poisson ΣΜΠ.

Πιο συγκεκριμένα, η Poisson ΣΜΠ με παράμετρο λ είναι μία καλή προσέγγιση της διωνυμικής ΣΜΠ με παραμέτρους n και p, δηλαδή,

$$e^{-\lambda}\frac{\lambda^k}{k!} \approx \frac{n!}{k!\,(n-k)!}p^k(1-p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

με την προϋπόθεση ότι $\lambda = np$, το n είναι πολύ μεγάλο και το p πολύ μικρό. Στην περίπτωση αυτή η χρήση της Poisson ΣΜΠ μπορεί να δώσει απλούστερα μοντέλα και υπολογισμούς. Έστω $n = 100$ και $p = 0.01$. Τότε η πιθανότητα $k = 5$ επιτυχιών στις $n = 100$ επαναλήψεις υπολογίζεται με τη χρήση της διωνυμικής ΣΜΠ ως εξής:

$$\frac{100!}{95!\,5!} \cdot 0.01^5(1-0.01)^{95} = 0.00290.$$

Χρησιμοποιώντας την Poisson με $\lambda = np = 100 \cdot 0.01 = 1$, η πιθανότητα αυτή προσεγγίζεται ως εξής:

$$e^{-1}\frac{1}{5!} = 0.00306.$$

Δίνουμε μία μαθηματική αιτιολόγηση της προσεγγιστικής ιδιότητας της Poisson στα προβλήματα στο τέλος του κεφαλαίου και επίσης στο Κεφάλαιο 5, όπου

θα την ερμηνεύσουμε περαιτέρω, θα την επεκτείνουμε και θα τη χρησιμοποιήσουμε για να περιγράψουμε τη διαδικασία Poisson.

2.3 ΣΥΝΑΡΤΗΣΕΙΣ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

Δεδομένης μίας τυχαίας μεταβλητής X, μπορούμε να ορίσουμε άλλες τυχαίες μεταβλητές εφαρμόζοντας διάφορους μετασχηματισμούς στην X. Για παράδειγμα, έστω ότι η τυχαία μεταβλητή X είναι η σημερινή θερμοκρασία σε βαθμούς Κελσίου και θεωρήστε το μετασχηματισμό $Y = 1.8X+32$, ο οποίος δίνει τη θερμοκρασία σε βαθμούς Φαρενάιτ. Στο παράδειγμα αυτό η Y είναι μία **γραμμική** συνάρτηση της X, της μορφής

$$Y = g(X) = aX + b,$$

όπου a και b είναι σταθερές. Επίσης μπορούμε να θεωρήσουμε και μη γραμμικές συναρτήσεις γενικής μορφής

$$Y = g(X).$$

Για παράδειγμα, εάν θέλουμε να παρουσιάσουμε τις θερμοκρασίες σε λογαριθμική κλίμακα, μπορούμε να χρησιμοποιήσουμε τη συνάρτηση $g(X) = \log X$.

Εάν $Y = g(X)$ είναι μία συνάρτηση μίας τυχαίας μεταβλητής X, τότε η Y είναι επίσης μία τυχαία μεταβλητή, επειδή δίνει μία αριθμητική τιμή για κάθε δυνατό αποτέλεσμα. Αυτό συμβαίνει επειδή κάθε αποτέλεσμα στο δειγματικό χώρο ορίζει μία αριθμητική τιμή x για την X και επομένως μία αριθμητική τιμή $y = g(x)$ για την Y. Εάν η X είναι διακριτή με ΣΜΠ την p_X, τότε η Y είναι επίσης διακριτή και η ΣΜΠ της p_Y μπορεί να υπολογιστεί χρησιμοποιώντας τη ΣΜΠ της X. Συγκεκριμένα, για να πάρουμε την $p_Y(y)$ για οποιοδήποτε y, προσθέτουμε τις πιθανότητες όλων των τιμών x για τις οποίες $g(x) = y$:

$$p_Y(y) = \sum_{\{x \mid g(x)=y\}} p_X(x).$$

Παράδειγμα 2.1. Έστω $Y = |X|$. Ας εφαρμόσουμε τον προηγούμενο τύπο για την ΣΜΠ p_Y στην περίπτωση όπου

$$p_X(x) = \begin{cases} 1/9, & \text{εάν } x \text{ είναι ακέραιος στο διάστημα } [-4, 4], \\ 0, & \text{αλλιώς·} \end{cases}$$

βλ. Σχ. 2.6 . Οι δυνατές τιμές της Y είναι $y = 0, 1, 2, 3, 4$. Για να υπολογίσουμε την

$p_Y(y)$ για κάποια δεδομένη τιμή y στο διάστημα αυτό, χρειάζεται να προσθέσουμε την $p_X(x)$ για όλες τις τιμές x που ικανοποιούν την $|x| = y$. Πιο συγκεκριμένα, υπάρχει μόνο μία τιμή της X η οποία αντιστοιχεί στο $y = 0$, δηλαδή η $x = 0$. Επομένως,

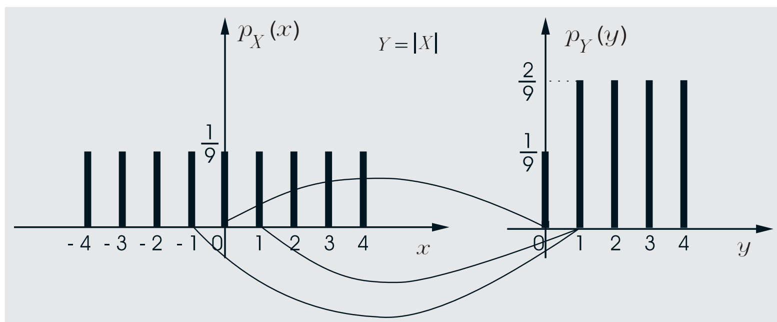
$$p_Y(0) = p_X(0) = \frac{1}{9}.$$

Επίσης, υπάρχουν δύο τιμές της X οι οποίες αντιστοιχούν σε κάθε $y = 1, 2, 3, 4$. Για παράδειγμα,

$$p_Y(1) = p_X(-1) + p_X(1) = \frac{2}{9}.$$

Επομένως, η ΣΜΠ της Y είναι

$$p_Y(y) = \begin{cases} 2/9, & \text{εάν } y = 1, 2, 3, 4, \\ 1/9, & \text{εάν } y = 0, \\ 0, & \text{αλλιώς.} \end{cases}$$



Σχήμα 2.6: Η ΣΜΠ της X και $Y = |X|$ στο Παράδειγμα 2.1.

Για ένα διαφορετικό παράδειγμα, έστω $Z = X^2$. Για να βρούμε τη ΣΜΠ της Z, μπορούμε να τη θεωρήσουμε ως το τετράγωνο της τυχαίας μεταβλητής X ή ως το τετράγωνο της τυχαίας μεταβλητής $Y = |X|$. Εφαρμόζοντας τον τύπο $p_Z(z) = \sum_{\{x \mid x^2 = z\}} p_X(x)$ ή τον τύπο $p_Z(z) = \sum_{\{y \mid y^2 = z\}} p_Y(y)$, έχουμε

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

2.4 ΜΕΣΗ ΤΙΜΗ ΚΑΙ ΔΙΑΣΠΟΡΑ

Η ΣΜΠ μίας τυχαίας μεταβλητής μας δίνει πολλούς αριθμούς, τις πιθανότητες όλων των δυνατών τιμών της X. Θα θέλαμε να συνοψίσουμε όλη αυτή την

πληροφορία σε ένα και μόνο αντιπροσωπευτικό αριθμό. Αυτό επιτυγχάνεται με τη **μέση τιμή** της X, η οποία είναι ο μέσος όρος των δυνατών τιμών της X σταθμισμένων ανάλογα με τις πιθανότητές τους.

Υποθέστε ότι περιστρέφετε ένα τροχό της τύχης πολλές φορές. Σε κάθε περιστροφή ένας από τους αριθμούς $m_1, m_2, \ldots, m_n$ προκύπτει με αντίστοιχες πιθανότητες $p_1, p_2, \ldots, p_n$ και αυτή είναι η αμοιβή σας για τη συγκεκριμένη περιστροφή. Ποιο είναι το χρηματικό ποσό που "αναμένετε" να πάρετε "ανά περιστροφή"; Οι όροι "αναμένετε" και "ανά περιστροφή" είναι ασαφείς, αλλά ας δούμε μία λογική ερμηνεία τους.

Υποθέστε ότι περιστρέφετε το τροχό k φορές και ότι k_i είναι ο αριθμός των φορών που το αποτέλεσμα είναι m_i. Τότε, το συνολικό ποσό που λαμβάνετε είναι $m_1k_1 + m_2k_2 + \cdots + m_nk_n$. Το συνολικό ποσό που λαμβάνετε ανά περιστροφή είναι

$$M = \frac{m_1k_1 + m_2k_2 + \cdots + m_nk_n}{k}.$$

Εάν ο αριθμός των περιστροφών k είναι πολύ μεγάλος και εάν είμαστε διατεθειμένοι να ερμηνεύσουμε τις πιθανότητες ως σχετικές συχνότητες, τότε είναι λογικό να υποθέσουμε ότι ο m_i θα εμφανισθεί με συχνότητα περίπου ίση με p_i:

$$\frac{k_i}{k} \approx p_i, \qquad i = 1, \ldots, n.$$

Άρα το ποσό των χρημάτων ανά περιστροφή που "αναμένετε" να λάβετε είναι

$$M = \frac{m_1k_1 + m_2k_2 + \cdots + m_nk_n}{k} \approx m_1p_1 + m_2p_2 + \cdots + m_np_n.$$

Με βάση το παραπάνω παράδειγμα, εισάγουμε τον ακόλουθο ορισμό.†

† Όταν έχουμε να κάνουμε με τυχαίες μεταβλητές, οι οποίες παίρνουν ένα αριθμήσιμα άπειρο πλήθος τιμών, υπάρχει περίπτωση το άπειρο άθροισμα $\sum_x xp_X(x)$ να μην είναι καλώς ορισμένο. Πιο συγκεκριμένα, θα λέμε ότι η μέση τιμή είναι καλώς ορισμένη εάν $\sum_x |x|p_X(x) < \infty$. Στην περίπτωση αυτή είναι γνωστό ότι το άπειρο άθροισμα $\sum_x xp_X(x)$ συγκλίνει σε μία πεπερασμένη τιμή, η οποία είναι ανεξάρτητη της σειράς με την οποία αθροίζονται οι διάφοροι όροι.

Δίνουμε ένα παράδειγμα όπου η μέση τιμή δεν είναι καλώς ορισμένη. Θεωρήστε μία τυχαία μεταβλητή X η οποία παίρνει την τιμή 2^k με πιθανότητα 2^{-k}, για $k = 1, 2, \ldots$. Για ένα πιο ενδιαφέρον παράδειγμα, θεωρήστε την τυχαία μεταβλητή X η οποία παίρνει τις τιμές 2^k και -2^k με πιθανότητα 2^{-k}, για $k = 2, 3, \ldots$. Η μέση τιμή ξανά δεν ορίζεται, αν και η ΣΜΠ είναι συμμετρική γύρω από το μηδέν και θα μπορούσε κανείς να θεωρήσει ότι η $\mathbf{E}[X]$ είναι μηδέν.

Στο υπόλοιπο του βιβλίου αυτού, εκτός εάν υπάρχει ένδειξη για το αντίθετο, υποθέτουμε ότι η μέση τιμή των τυχαίων μεταβλητών που μας ενδιαφέρουν είναι καλώς ορισμένη.

Μέση Τιμή

Ορίζουμε τη **μέση τιμή** μίας τυχαίας μεταβλητής X με ΣΜΠ p_X, ως εξής:

$$\mathbf{E}[X] = \sum_x x p_X(x).$$

Παράδειγμα 2.2. Θεωρήστε δύο ανεξάρτητες ρίψεις ενός νομίσματος, κάθε μία με πιθανότητα κορώνας ίση με 3/4 και έστω X ο αριθμός των κορωνών που παίρνουμε. Τότε, η X είναι μία διωνυμική τυχαία μεταβλητή με παραμέτρους $n = 2$ και $p = 3/4$. Η ΣΜΠ της είναι

$$p_X(k) = \begin{cases} (1/4)^2, & \text{εάν } k = 0, \\ 2 \cdot (1/4) \cdot (3/4), & \text{εάν } k = 1, \\ (3/4)^2, & \text{εάν } k = 2, \end{cases}$$

άρα η μέση τιμή είναι

$$\mathbf{E}[X] = 0 \cdot \left(\frac{1}{4}\right)^2 + 1 \cdot \left(2 \cdot \frac{1}{4} \cdot \frac{3}{4}\right) + 2 \cdot \left(\frac{3}{4}\right)^2 = \frac{24}{16} = \frac{3}{2}.$$

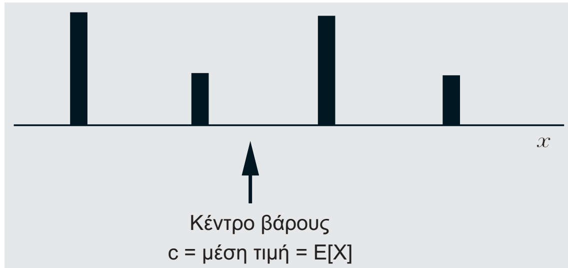
Είναι χρήσιμο να θεωρούμε τη μέση τιμή της X ως μία “αντιπροσωπευτική” τιμή της X, η οποία βρίσκεται κάπου στο μέσο των δυνατών τιμών της. Μπορεί να είμαστε πιο ακριβείς θεωρώντας τη μέση τιμή ως το **κέντρο βάρους** της ΣΜΠ, υπό την έννοια που επεξηγείται στο Σχ. 2.7.

Διασπορά, Ροπές και ο Κανόνας της Μέσης Τιμής

Εκτός από τη μέση τιμή, υπάρχουν πολλές άλλες ποσότητες τις οποίες μπορούμε να συνδέσουμε με μία τυχαία μεταβλητή και τη ΣΜΠ της. Για παράδειγμα, ορίζουμε τη 2**η ροπή** της τυχαίας μεταβλητής X ως τη μέση τιμή της τυχαίας μεταβλητής X^2. Πιο γενικά ορίζουμε την n-**οστή ροπή** $\mathbf{E}[X^n]$ ως τη μέση τιμή της τυχαίας μεταβλητής X^n. Σύμφωνα με την ορολογία αυτή η 1η ροπή της X είναι απλώς η μέση τιμή.

Η πιο σημαντική ποσότητα που σχετίζεται με μία τυχαία μεταβλητή X, εκτός της μέσης τιμής, είναι η **διασπορά**, η οποία δηλώνεται με var(X) και ορίζεται ως η μέση τιμή της τυχαίας μεταβλητής $(X - \mathbf{E}[X])^2$, δηλαδή,

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right].$$



Σχήμα 2.7: Ερμηνεία της μέσης τιμής ως κέντρο βάρους. Δεδομένης μίας μάζας με βάρος $p_X(x)$ η οποία τοποθετείται σε κάθε σημείο x με $p_X(x) > 0$, το κέντρο βάρους c είναι το σημείο στο οποίο το άθροισμα των ροπών από τα βάρη αριστερά είναι ίσο με το άθροισμα των ροπών από τα βάρη δεξιά:

$$\sum_x (x-c)p_X(x) = 0.$$

Άρα $c = \sum_x x p_X(x)$, δηλαδή το κέντρο βάρους είναι ίσο με τη μέση τιμή $\mathbf{E}[X]$.

Επειδή η $(X - \mathbf{E}[X])^2$ παίρνει μη αρνητικές τιμές, η διασπορά είναι μη αρνητική.

Η διασπορά είναι ένα μέτρο της διακύμανσης των τιμών της X γύρω από τη μέση τιμή. Ένα άλλο μέτρο διακύμανσης είναι η **τυπική απόκλιση** της X, η οποία ορίζεται ως η τετραγωνική ρίζα της διασποράς και συμβολίζεται με σ_X:

$$\sigma_X = \sqrt{\text{var}(X)}.$$

Η τυπική απόκλιση είναι πιο εύκολο να ερμηνευτεί γιατί έχει τις ίδιες μονάδες όπως η X. Για παράδειγμα, εάν η X μετράει μήκος σε μέτρα, οι μονάδες της διασποράς είναι τετραγωνικά μέτρα, ενώ οι μονάδες της τυπικής απόκλισης είναι μέτρα.

Ένας τρόπος να υπολογιστεί η διασπορά var(X) είναι με τη χρήση του ορισμού της μέσης τιμής, αφού υπολογιστεί η ΣΜΠ της τυχαίας μεταβλητής $(X - \mathbf{E}[X])^2$. Αυτή η τυχαία μεταβλητή είναι μία συνάρτηση της X, και η ΣΜΠ της υπολογίζεται με τον τρόπο που συζητήθηκε στην προηγούμενη παράγραφο.

Παράδειγμα 2.3. Θεωρήστε την τυχαία μεταβλητή X του Παραδείγματος 2.1, η οποία έχει ΣΜΠ

$$p_X(x) = \begin{cases} 1/9, & \text{εάν } x \text{ είναι ακέραιος στο διάστημα } [-4,4], \\ 0, & \text{αλλιώς.} \end{cases}$$

Η μέση τιμή $\mathbf{E}[X]$ είναι ίση με 0. Αυτό φαίνεται από τη συμμετρία της ΣΜΠ της X γύρω από το 0 και επαληθεύεται από τον ορισμό:

$$\mathbf{E}[X] = \sum_x x p_X(x) = \frac{1}{9} \sum_{x=-4}^{4} x = 0.$$

Έστω $Z = \big(X - \mathbf{E}[X]\big)^2 = X^2$. Όπως στο Παράδειγμα 2.1, έχουμε

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

Η διασπορά της X είναι

$$\text{var}(X) = \mathbf{E}[Z] = \sum_z z p_Z(z) = 0 \cdot \frac{1}{9} + 1 \cdot \frac{2}{9} + 4 \cdot \frac{2}{9} + 9 \cdot \frac{2}{9} + 16 \cdot \frac{2}{9} = \frac{60}{9}.$$

Υπάρχει μία ευκολότερη μέθοδος να υπολογίσουμε τη διασπορά var(X), ο οποίος χρησιμοποιεί τη ΣΜΠ της X αλλά **δεν προϋποθέτει τη ΣΜΠ της** $(X - \mathbf{E}[X])^2$. Η μέθοδος αυτή βασίζεται στον παρακάτω κανόνα.

Κανόνας της Μέσης Τιμής για Συναρτήσεις Τυχαίων Μεταβλητών

Έστω X μία τυχαία μεταβλητή με ΣΜΠ p_X και έστω $g(X)$ μία συνάρτηση της X. Τότε η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x).$$

Για να επαληθεύσουμε τον κανόνα αυτό, θέτουμε $Y = g(X)$ και χρησιμοποιούμε τον τύπο

$$p_Y(y) = \sum_{\{x \mid g(x)=y\}} p_X(x)$$

τον οποίο έχουμε αποδείξει στην προηγούμενη παράγραφο. Έχουμε

$$\begin{aligned} \mathbf{E}[g(X)] &= \mathbf{E}[Y] \\ &= \sum_y y p_Y(y) \end{aligned}$$

$$
\begin{aligned}
&= \sum_y y \sum_{\{x \,|\, g(x)=y\}} p_X(x) \\
&= \sum_y \sum_{\{x \,|\, g(x)=y\}} y p_X(x) \\
&= \sum_y \sum_{\{x \,|\, g(x)=y\}} g(x) p_X(x) \\
&= \sum_x g(x) p_X(x).
\end{aligned}
$$

Χρησιμοποιώντας τον κανόνα της μέσης τιμής, μπορούμε να γράψουμε τη διασπορά της X ως

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] = \sum_x (x - \mathbf{E}[X])^2 p_X(x).$$

Παρόμοια, η n-οστή ροπή είναι

$$\mathbf{E}[X^n] = \sum_x x^n p_X(x),$$

και δεν χρειάζεται να υπολογιστεί η ΣΜΠ της X^n.

Παράδειγμα 2.3 (συνέχεια). Για την τυχαία μεταβλητή X με ΣΜΠ

$$p_X(x) = \begin{cases} 1/9, & \text{εάν } x \text{ είναι ένας ακέραιος στο διάστημα } [-4, 4], \\ 0, & \text{αλλιώς,} \end{cases}$$

έχουμε

$$
\begin{aligned}
\text{var}(X) &= \mathbf{E}\left[\big(X - \mathbf{E}[X]\big)^2\right] \\
&= \sum_x \big(x - \mathbf{E}[X]\big)^2 p_X(x) \\
&= \frac{1}{9} \sum_{x=-4}^{4} x^2 \qquad\qquad (\text{επειδή } \mathbf{E}[X] = 0) \\
&= \frac{1}{9}(16 + 9 + 4 + 1 + 0 + 1 + 4 + 9 + 16) \\
&= \frac{60}{9},
\end{aligned}
$$

το οποίο συμφωνεί με το αποτέλεσμα που βρήκαμε νωρίτερα.

Σημειώσαμε νωρίτερα ότι η διασπορά είναι πάντα μη αρνητική. Θα μπορούσε όμως να είναι μηδέν; Καθώς κάθε όρος στον τύπο $\sum_x (x - \mathbf{E}[X])^2 p_X(x)$ της διασποράς είναι μη αρνητικός, το άθροισμα είναι μηδέν εάν και μόνο εάν $(x - \mathbf{E}[X])^2 p_X(x) = 0$ για κάθε x. Αυτή η συνθήκη συνεπάγεται ότι για κάθε x με $p_X(x) > 0$, πρέπει να έχουμε $x = \mathbf{E}[X]$ ώστε η τυχαία μεταβλητή X να μην είναι πραγματικά "τυχαία": η τιμή της είναι ίση με τη μέση τιμή $\mathbf{E}[X]$, με πιθανότητα 1.

Διασπορά

Η διασπορά $\text{var}(X)$ μίας τυχαίας μεταβλητής X ορίζεται ως

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right],$$

και μπορεί να υπολογιστεί ως εξής:

$$\text{var}(X) = \sum_x (x - \mathbf{E}[X])^2 p_X(x).$$

Είναι πάντα μη αρνητική. Η τετραγωνική της ρίζα δηλώνεται με σ_X και ονομάζεται **τυπική απόκλιση**.

Ιδιότητες της Μέσης Τιμής και Διασποράς

Θα χρησιμοποιήσουμε τώρα τον κανόνα της μέσης τιμής για να αποδείξουμε μερικές σημαντικές ιδιότητες της μέσης τιμής και της διασποράς. Αρχίζουμε με μία τυχαία μεταβλητή X και ορίζουμε μία νέα τυχαία μεταβλητή Y, της μορφής

$$Y = aX + b,$$

όπου a και b είναι γνωστές σταθερές. Θα υπολογίσουμε τη μέση τιμή και τη διασπορά της γραμμικής συνάρτησης Y. Έχουμε

$$\mathbf{E}[Y] = \sum_x (ax + b) p_X(x) = a \sum_x x p_X(x) + b \sum_x p_X(x) = a\mathbf{E}[X] + b.$$

Επιπλέον,

$$\text{var}(Y) = \sum_x (ax + b - \mathbf{E}[aX + b])^2 p_X(x)$$

$$
\begin{aligned}
&= \sum_x (ax + b - a\mathbf{E}[X] - b)^2 p_X(x) \\
&= a^2 \sum_x (x - \mathbf{E}[X])^2 p_X(x) \\
&= a^2 \operatorname{var}(X).
\end{aligned}
$$

Μέση Τιμή και Διασπορά Γραμμικής Συνάρτησης μίας Τυχαίας Μεταβλητής

Έστω X μία τυχαία μεταβλητή και έστω

$$Y = aX + b,$$

όπου a και b είναι γνωστές σταθερές. Τότε,

$$\mathbf{E}[Y] = a\mathbf{E}[X] + b, \qquad \operatorname{var}(Y) = a^2 \operatorname{var}(X).$$

Ας δώσουμε τώρα ένα χρήσιμο τύπο για τη διασπορά μίας τυχαίας μεταβλητής X με δεδομένη ΣΜΠ.

Σχέση Διασποράς και Ροπών

$$\operatorname{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2.$$

Η έκφραση αυτή επαληθεύεται ως ακολούθως:

$$
\begin{aligned}
\operatorname{var}(X) &= \sum_x (x - \mathbf{E}[X])^2 p_X(x) \\
&= \sum_x \Big(x^2 - 2x\mathbf{E}[X] + (\mathbf{E}[X])^2\Big) p_X(x) \\
&= \sum_x x^2 p_X(x) - 2\mathbf{E}[X] \sum_x x p_X(x) + (\mathbf{E}[X])^2 \sum_x p_X(x) \\
&= \mathbf{E}[X^2] - 2(\mathbf{E}[X])^2 + (\mathbf{E}[X])^2 \\
&= \mathbf{E}[X^2] - (\mathbf{E}[X])^2.
\end{aligned}
$$

Τέλος, δείχνουμε με ένα παράδειγμα ένα λάθος που γίνεται συχνά: γενικά η $\mathbf{E}[g(X)]$ δεν είναι ίση με τη $g(\mathbf{E}[X])$, εκτός εάν η $g(X)$ είναι μία γραμμική συνάρτηση.

Παράδειγμα 2.4. Μέση Ταχύτητα και Μέσος Χρόνος. Εάν ο καιρός είναι καλός (το οποίο συμβαίνει με πιθανότητα 0.6), η Αλίκη περπατάει δύο χιλιόμετρα για να φτάσει στην τάξη της με ταχύτητα $V = 5$ χιλιόμετρα την ώρα, και διαφορετικά παίρνει την μοτοσυκλέτα της και φτάνει στην τάξη της με ταχύτητα $V = 30$ χιλιόμετρα την ώρα. Ποια είναι η μέση τιμή του χρόνου T για να φτάσει στην τάξη της;

Ο σωστός τρόπος για να λύσουμε το πρόβλημα είναι να βρούμε πρώτα τη ΣΜΠ της T,

$$p_T(t) = \begin{cases} 0.6, & \text{εάν } t = 2/5 \text{ ώρες}, \\ 0.4, & \text{εάν } t = 2/30 \text{ ώρες}, \end{cases}$$

και μετά να υπολογίσουμε τη μέση τιμή

$$\mathbf{E}[T] = 0.6 \cdot \frac{2}{5} + 0.4 \cdot \frac{2}{30} = \frac{4}{15} \text{ ώρες}.$$

Θα ήταν λάθος να βρούμε τη μέση τιμή της ταχύτητας V,

$$\mathbf{E}[V] = 0.6 \cdot 5 + 0.4 \cdot 30 = 15 \text{ χιλιόμετρα την ώρα},$$

και κατόπιν να ισχυριστούμε ότι η μέση τιμή της T είναι

$$\frac{2}{\mathbf{E}[V]} = \frac{2}{15} \text{ ώρες}.$$

Συνοψίζοντας, σ' αυτό το παράδειγμα έχουμε

$$T = \frac{2}{V}, \qquad \text{και } \mathbf{E}[T] = \mathbf{E}\left[\frac{2}{V}\right] \neq \frac{2}{\mathbf{E}[V]}.$$

Μέσες Τιμές και Διασπορές Μερικών Κοινών Τυχαίων Μεταβλητών

Θα αποδείξουμε τώρα τους τύπους της μέσης τιμής και διασποράς μερικών σημαντικών τυχαίων μεταβλητών. Οι τύποι αυτοί θα χρησιμοποιηθούν επανειλημμένα στο υπόλοιπο του βιβλίου.

Παράδειγμα 2.5. Μέση Τιμή και Διασπορά της Bernoulli. Θεωρήστε το πείραμα ρίψης ενός νομίσματος στο οποίο εμφανίζεται κορώνα με πιθανότητα p και γράμματα

με πιθανότητα $1-p$, και την Bernoulli τυχαία μεταβλητή X με ΣΜΠ

$$p_X(k) = \begin{cases} p, & \text{εάν } k = 1, \\ 1-p, & \text{εάν } k = 0. \end{cases}$$

Η μέση τιμή, η δεύτερη ροπή και η διασπορά της υπολογίζονται ως εξής:

$$\begin{aligned} \mathbf{E}[X] &= 1 \cdot p + 0 \cdot (1-p) = p, \\ \mathbf{E}[X^2] &= 1^2 \cdot p + 0 \cdot (1-p) = p, \\ \text{var}(X) &= \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2 = p - p^2 = p(1-p). \end{aligned}$$

Παράδειγμα 2.6. Διακριτή Ομοιόμορφη Τυχαία Μεταβλητή. Ποια είναι η μέση τιμή και η διασπορά της ρίψης ενός αμερόληπτου 6-εδρου ζαριού; Εάν θεωρήσουμε το αποτέλεσμα της ρίψης ως μία τυχαία μεταβλητή X, η ΣΜΠ της είναι

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

Επειδή η ΣΜΠ είναι συμμετρική γύρω από το 3.5, έχουμε $\mathbf{E}[X] = 3.5$. Για ότι αφορά τη διασπορά έχουμε

$$\begin{aligned} \text{var}(X) &= \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2 \\ &= \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - (3.5)^2, \end{aligned}$$

και παίρνουμε $\text{var}(X) = 35/12$.

Η παραπάνω τυχαία μεταβλητή είναι ειδική περίπτωση μίας **διακριτής ομοιόμορφα κατανεμημένης** τυχαίας μεταβλητής (ή **διακριτής ομοιόμορφης** για συντομία), η οποία εξ ορισμού παίρνει μία οποιαδήποτε τιμή σε ένα πεδίο διαδοχικών ακέραιων τιμών, με ίση πιθανότητα. Πιο συγκεκριμένα, η τυχαία μεταβλητή αυτή έχει μία ΣΜΠ της μορφής

$$p_X(k) = \begin{cases} \dfrac{1}{b-a+1}, & \text{εάν } k = a, a+1, \ldots, b, \\ 0, & \text{αλλιώς,} \end{cases}$$

όπου a και b είναι δύο ακέραιοι με $a < b$· βλ. Σχ. 2.8.

Η μέση τιμή είναι

$$\mathbf{E}[X] = \frac{a+b}{2},$$

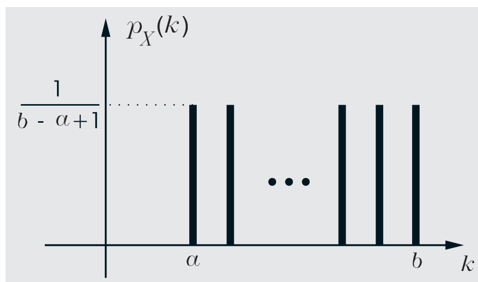
όπως διαπιστώνεται από την παρατήρηση ότι η ΣΜΠ είναι συμμετρική γύρω από το $(a+b)/2$. Για να υπολογίσουμε τη διασπορά της X, πρώτα θεωρούμε την πιο απλή

περίπτωση όπου $a = 1$ και $b = n$. Επαληθεύεται επαγωγικά ως προς n ότι

$$\mathbf{E}[X^2] = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{6}(n+1)(2n+1).$$

Αφήνουμε την επαλήθευση αυτή σαν άσκηση για τον αναγνώστη. Η διασπορά τώρα προκύπτει ως συνάρτηση της πρώτης και δεύτερης ροπής.

$$\begin{aligned}\mathrm{var}(X) &= \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2 \\ &= \frac{1}{6}(n+1)(2n+1) - \frac{1}{4}(n+1)^2 \\ &= \frac{1}{12}(n+1)(4n+2-3n-3) \\ &= \frac{n^2-1}{12}.\end{aligned}$$



Σχήμα 2.8: Η ΣΜΠ μίας διακριτής τυχαίας μεταβλητής η οποία είναι ομοιόμορφα κατανεμημένη μεταξύ δύο ακεραίων a και b. Η μέση τιμή και η διασπορά είναι

$$\mathbf{E}[X] = \frac{a+b}{2}, \qquad \mathrm{var}(X) = \frac{(b-a)(b-a+2)}{12}.$$

Για την περίπτωση γενικών ακέραιων αριθμών a και b, σημειώνουμε ότι μία τυχαία μεταβλητή η οποία είναι ομοιόμορφα κατανεμημένη στο διάστημα $[a, b]$ έχει την ίδια διασπορά όπως μία τυχαία μεταβλητή η οποία είναι ομοιόμορφα κατανεμημένη στο διάστημα $[1, b-a+1]$, επειδή η ΣΜΠ της δεύτερης είναι απλώς η μετατοπισμένη ΣΜΠ της πρώτης. Άρα η ζητούμενη διασπορά προκύπτει από τον παραπάνω τύπο θέτοντας $n = b - a + 1$, ως εξής

$$\mathrm{var}(X) = \frac{(b-a+1)^2 - 1}{12} = \frac{(b-a)(b-a+2)}{12}.$$

Παράδειγμα 2.7. Η Μέση Τιμή της Poisson. Η μέση τιμή της Poisson ΣΜΠ

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots,$$

μπορεί να υπολογιστεί όπως φαίνεται παρακάτω:

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} && (\text{ο } k = 0 \text{ όρος δίνει μηδέν}) \\
&= \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda \sum_{m=0}^{\infty} e^{-\lambda} \frac{\lambda^m}{m!} && (\text{θέτοντας } m = k-1) \\
&= \lambda.
\end{aligned}$$

Η τελευταία ισότητα ισχύει διότι

$$\sum_{m=0}^{\infty} e^{-\lambda} \frac{\lambda^m}{m!} = \sum_{m=0}^{\infty} p_X(m) = 1$$

που είναι η ιδιότητα της κανονικοποίησης της Poisson ΣΜΠ.

Ένας παρόμοιος υπολογισμός δίνει τη διασπορά της τυχαίας μεταβλητής Poisson, που είναι επίσης ίση με λ (βλ. προβλήματα στο τέλος του κεφαλαίου). Θα αποδείξουμε το γεγονός αυτό με διαφορετικούς τρόπους στα επόμενα κεφάλαια.

Λήψη Αποφάσεων Χρησιμοποιώντας Μέσες Τιμές

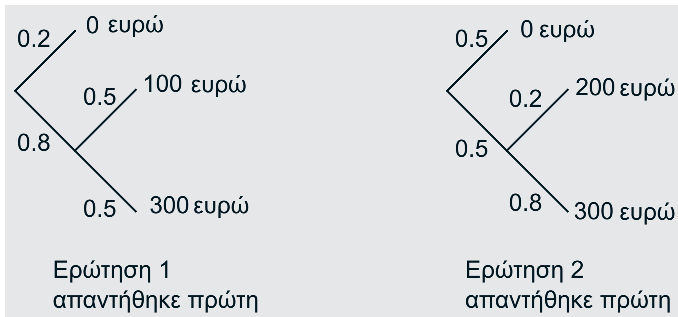
Οι μέσες τιμές συχνά μας διευκολύνουν να διατυπώσουμε προβλήματα βελτιστοποίησης, όπου έχουμε να επιλέξουμε μεταξύ πιθανών αποφάσεων που αποφέρουν τυχαία κέρδη. Εάν θεωρήσουμε το αναμενόμενο κέρδος μίας απόφασης ως το "μέσο κέρδος για ένα μεγάλο αριθμό επαναλήψεων", είναι λογικό να επιλέξουμε μία απόφαση με μέγιστο αναμενόμενο κέρδος. Δίνουμε ένα τέτοιο παράδειγμα παρακάτω.

Παράδειγμα 2.8. Το Πρόβλημα του Παιχνιδιού Γνώσεων. Το παράδειγμα αυτό, όταν γενικευτεί κατάλληλα, αποτελεί πρότυπο μεθόδου βέλτιστης δρομολόγησης διεργασιών που έχουν τυχαία αποτελέσματα.

Θεωρήστε ένα παιχνίδι γνώσεων, όπου κάποιο άτομο πρέπει να αποφασίσει τη σειρά με την οποία θα απαντήσει δύο ερωτήσεις. Η ερώτηση 1 θα απαντηθεί σωστά με πιθανότητα 0.8 και τότε το άτομο θα λάβει σαν βραβείο 100 ευρώ, ενώ η ερώτηση 2 θα απαντηθεί σωστά με πιθανότητα 0.5 και το άτομο θα λάβει σαν βραβείο 200 ευρώ. Εάν η πρώτη απάντηση είναι λάθος το παιχνίδι τελειώνει, δηλαδή δεν επιτρέπεται στο άτομο να επιχειρήσει τη δεύτερη ερώτηση. Εάν η πρώτη απάντηση είναι σωστή, τότε

επιτρέπεται στο άτομο να επιχειρήσει τη δεύτερη ερώτηση. Ποια ερώτηση θα πρέπει να απαντηθεί πρώτα για να μεγιστοποιηθεί η μέση τιμή του συνολικού χρηματικού βραβείου;

Η απάντηση δεν είναι προφανής διότι αν επιχειρήσει κανείς να απαντήσει πρώτα την πιο προσοδοφόρα αλλά συγχρόνως και πιο δύσκολη ερώτηση 2, κινδυνεύει να χάσει την ευκαιρία να απαντήσει την ευκολότερη ερώτηση 1. Ας θεωρήσουμε το συνολικό βραβείο ως τυχαία μεταβλητή X και ας υπολογίσουμε τη μέση τιμή $\mathbf{E}[X]$ για τις δύο διαφορετικές σειρές των ερωτήσεων (βλ. Σχ. 2.9):



Σχήμα 2.9: Ακολουθιακή περιγραφή του δειγματικού χώρου του παιχνιδιού γνώσεων για τις περιπτώσεις που απαντάται πρώτα η ερώτηση 1 ή πρώτα η ερώτηση 2.

(α) *Απαντάται πρώτα η ερώτηση 1*: Τότε η ΣΜΠ της X είναι (βλ. την αριστερή πλευρά του Σχ. 2.9)

$$p_X(0) = 0.2, \qquad p_X(100) = 0.8 \cdot 0.5, \qquad p_X(300) = 0.8 \cdot 0.5,$$

και έχουμε

$$\mathbf{E}[X] = 0.8 \cdot 0.5 \cdot 100 + 0.8 \cdot 0.5 \cdot 300 = 160.$$

(β) *Απαντάται πρώτα η ερώτηση 2*: Τότε η ΣΜΠ της X είναι (βλ. την δεξιά πλευρά του Σχ. 2.9)

$$p_X(0) = 0.5, \qquad p_X(200) = 0.5 \cdot 0.2, \qquad p_X(300) = 0.5 \cdot 0.8,$$

και έχουμε

$$\mathbf{E}[X] = 0.5 \cdot 0.2 \cdot 200 + 0.5 \cdot 0.8 \cdot 300 = 140.$$

Άρα είναι προτιμότερο να επιχειρήσουμε να απαντήσουμε πρώτα την πιο εύκολη ερώτηση.

Ας γενικεύσουμε την ανάλυση. Δηλώνουμε με p_1 και p_2 τις πιθανότητες να απαντήσουμε σωστά την ερώτηση 1 και την ερώτηση 2, και με v_1 και v_2 τα αντίστοιχα βραβεία. Εάν η ερώτηση 1 απαντηθεί πρώτη, έχουμε

$$\mathbf{E}[X] = p_1(1-p_2)v_1 + p_1p_2(v_1+v_2) = p_1v_1 + p_1p_2v_2,$$

ενώ εάν η ερώτηση 2 απαντηθεί πρώτη, έχουμε

$$\mathbf{E}[X] = p_2(1-p_1)v_2 + p_2p_1(v_2+v_1) = p_2v_2 + p_2p_1v_1.$$

Άρα είναι βέλτιστο να απαντήσουμε πρώτα την ερώτηση 1 εάν και μόνο εάν

$$p_1v_1 + p_1p_2v_2 \geq p_2v_2 + p_2p_1v_1,$$

ή ισοδύναμα

$$\frac{p_1v_1}{1-p_1} \geq \frac{p_2v_2}{1-p_2}.$$

Επομένως η έκφραση $pv/(1-p)$ αποτελεί ένα αντιπροσωπευτικό δείκτη ποιότητας που καθορίζει τη βέλτιστη σειρά απαντήσεως των ερωτήσεων. Είναι ενδιαφέρον ότι αυτός ο κανόνας γενικεύεται στην περίπτωση που οι ερωτήσεις είναι περισσότερες από δύο (βλ. τα προβλήματα στο τέλος του κεφαλαίου).

2.5 ΑΠΟ ΚΟΙΝΟΥ ΣΜΠ ΠΟΛΛΑΠΛΩΝ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

Τα μοντέλα πιθανοτήτων συχνά αφορούν πολλές τυχαίες μεταβλητές. Για παράδειγμα, στην περίπτωση μίας ιατρικής διάγνωσης τα αποτελέσματα πολλαπλών τέστ ενδέχεται να είναι σημαντικά, ή στην περίπτωση δικτύων υπολογιστών, είναι δυνατόν να μας ενδιαφέρουν τα φορτία πολλών πυλών. Όλες αυτές οι τυχαίες μεταβλητές συνδέονται με το ίδιο πείραμα δειγματικό χώρο και νόμο πιθανότητας, και οι τιμές τους μπορεί να σχετίζονται με ενδιαφέροντες τρόπους. Αυτό μας οδηγεί στο να θεωρήσουμε πιθανότητες που συνδέουν συγχρόνως πολλές τυχαίες μεταβλητές, και να ερευνήσουμε την εφαρμογή τους. Στην παράγραφο αυτή θα επεκτείνουμε τις έννοιες της ΣΜΠ και της μέσης τιμής που έχουμε αναπτύξει μέχρι τώρα σε πολλαπλές τυχαίες μεταβλητές. Αργότερα, θα αναπτύξουμε έννοιες δέσμευσης και ανεξαρτησίας οι οποίες είναι παράλληλες με τις ιδέες που αναπτύξαμε στο Κεφάλαιο 1.

Θεωρήστε δύο διακριτές τυχαίες μεταβλητές X και Y που συνδέονται με το ίδιο πείραμα. Οι πιθανότητες των τιμών των X και Y εκφράζονται από την **από κοινού ΣΜΠ** των X και Y, η οποία δηλώνεται με $p_{X,Y}$. Συγκεκριμένα, εάν (x, y) είναι ένα ζεύγος δυνατών τιμών των X και Y, η μάζα πιθανότητας του (x, y) είναι η πιθανότητα του γεγονότος $\{X = x,\, Y = y\}$:

$$p_{X,Y}(x, y) = \mathrm{P}(X = x, Y = y).$$

Εδώ χρησιμοποιούμε το συντομευμένο συμβολισμό $\mathrm{P}(X = x, Y = y)$ αντί για τον πιο ακριβή συμβολισμό $\mathrm{P}(\{X = x\} \cap \{Y = y\})$ ή $\mathrm{P}(X = x$ και $Y = x)$.

Η από κοινού ΣΜΠ δίνει την πιθανότητα οποιουδήποτε γεγονότος που είναι σχετικό με τις τυχαίες μεταβλητές X και Y. Για παράδειγμα, εάν A είναι το σύνολο όλων των ζευγών (x, y) τα οποία έχουν μία συγκεκριμένη ιδιότητα, τότε

$$\mathrm{P}((X, Y) \in A) = \sum_{(x,y) \in A} p_{X,Y}(x, y).$$

Μάλιστα, μπορούμε να υπολογίσουμε τις ΣΜΠ των X και Y χρησιμοποιώντας τους τύπους

$$p_X(x) = \sum_y p_{X,Y}(x, y), \qquad p_Y(y) = \sum_x p_{X,Y}(x, y).$$

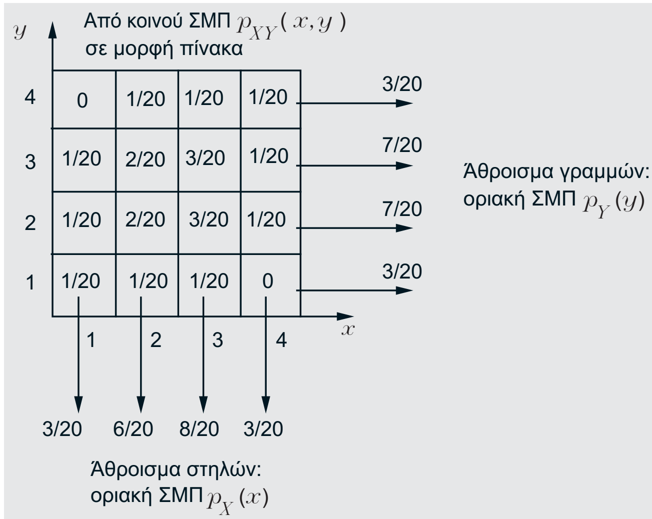
Ο τύπος για την $p_X(x)$ επαληθεύεται χρησιμοποιώντας τον υπολογισμό

$$\begin{aligned} p_X(x) &= \mathrm{P}(X = x) \\ &= \sum_y \mathrm{P}(X = x, Y = y) \\ &= \sum_y p_{X,Y}(x, y), \end{aligned}$$

όπου η δεύτερη ισότητα προκύπτει από την παρατήρηση ότι το γεγονός $\{X = x\}$ είναι η ένωση των ξένων μεταξύ τους γεγονότων $\{X = x, Y = y\}$ όταν το y παίρνει όλες τις διαφορετικές τιμές της Y. Ο τύπος $p_Y(y)$ επαληθεύεται παρόμοια. Θα αναφερόμαστε στις p_X και p_Y ως **οριακές** ΣΜΠ, για να τις ξεχωρίσουμε από τις από κοινού ΣΜΠ.

Μπορούμε να υπολογίσουμε οριακές ΣΜΠ από την από κοινού ΣΜΠ χρησιμοποιώντας την **μέθοδο του πίνακα**. Εδώ, η από κοινού ΣΜΠ των X και Y καταχωρείται σε διδιάστατο πίνακα και η **οριακή ΣΜΠ της X ή Y για κάποια δεδομένη τιμή λαμβάνεται προσθέτοντας τους αριθμούς κατά μήκος των αντίστοιχων στηλών ή γραμμών του πίνακα.** Βλέπε το παρακάτω Παράδειγμα και το Σχ. 2.10.

Παράδειγμα 2.9. Θεωρήστε δύο τυχαίες μεταβλητές X και Y οι οποίες περιγράφονται από την από κοινού ΣΜΠ όπως δείχνεται στο Σχ. 2.10. Οι οριακές ΣΜΠ υπολογίζονται προσθέτοντας τις καταχωρήσεις κατά μήκος των στηλών (για την οριακή ΣΜΠ της X) και κατά μήκος των γραμμών (για την ΣΜΠ της Y).



Σχήμα 2.10: Αναπαράσταση της μεθόδου του πίνακα για τον υπολογισμό των οριακών ΣΜΠ χρησιμοποιώντας την από κοινού ΣΜΠ του Παραδείγματος 2.9. Η από κοινού ΣΜΠ καταχωρείται στον πίνακα, όπου ο αριθμός σε κάθε τετράγωνο (x, y) δίνει την τιμή της $p_{X,Y}(x, y)$. Για να υπολογιστεί η οριακή ΣΜΠ $p_X(x)$ για μία δεδομένη τιμή x, προσθέτουμε τους αριθμούς της στήλης που αντιστοιχεί στην τιμή x. Για παράδειγμα $p_X(2) = 6/20$. Παρόμοια, για να υπολογιστεί η οριακή ΣΜΠ $p_Y(y)$ για κάποια δεδομένη τιμή y, προσθέτουμε τους αριθμούς της γραμμής που αντιστοιχεί στην τιμή y. Για παράδειγμα $p_Y(2) = 7/20$.

Συναρτήσεις Πολλαπλών Τυχαίων Μεταβλητών

Όταν υπάρχουν πολλαπλές τυχαίες μεταβλητές που μας ενδιαφέρουν, είναι δυνατό να δημιουργήσουμε καινούργιες τυχαίες μεταβλητές, θεωρώντας συναρτήσεις που συνδέουν κάποιες από αυτές. Πιο συγκεκριμένα, η συνάρτηση $Z = g(X, Y)$ των τυχαίων μεταβλητών X και Y ορίζει μία νέα τυχαία μεταβλητή. Η ΣΜΠ της μπορεί να υπολογιστεί από την από κοινού ΣΜΠ $p_{X,Y}$ ως εξής:

$$p_Z(z) = \sum_{\{(x,y)\,|\,g(x,y)=z\}} p_{X,Y}(x, y).$$

Επιπλέον, ο κανόνας της μέσης τιμής για συναρτήσεις επεκτείνεται και διαμορφώνεται ως

$$\mathbf{E}[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y).$$

Η επαλήθευση στην περίπτωση αυτή είναι παρόμοια με την προηγούμενη που αφορούσε συνάρτηση μίας τυχαίας μεταβλητής. Στην ειδική περίπτωση όπου η συνάρτηση g είναι γραμμική της μορφής $aX + bY + c$, όπου a, b και c είναι

γνωστές σταθερές, έχουμε

$$\mathbf{E}[aX + bY + c] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c.$$

Παράδειγμα 2.9 (συνέχεια). Θεωρήστε ξανά τις τυχαίες μεταβλητές X και Y, και την από κοινού ΣΜΠ τους όπως δίνεται στο Σχ. 2.10, καθώς και μία καινούργια τυχαία μεταβλητή Z η οποία ορίζεται ως

$$Z = X + 2Y.$$

Η ΣΜΠ της Z μπορεί να υπολογιστεί με τον τύπο

$$p_Z(z) = \sum_{\{(x,y)\,|\,x+2y=z\}} p_{X,Y}(x,y),$$

και χρησιμοποιώντας τη ΣΜΠ η οποία δίνεται στο Σχ. 2.10, έχουμε

$$p_Z(3) = \frac{1}{20},\ \ p_Z(4) = \frac{1}{20},\ \ p_Z(5) = \frac{2}{20},\ \ p_Z(6) = \frac{2}{20},\ \ p_Z(7) = \frac{4}{20},$$

$$p_Z(8) = \frac{3}{20},\ \ p_Z(9) = \frac{3}{20},\ \ p_Z(10) = \frac{2}{20},\ \ p_Z(11) = \frac{1}{20},\ \ p_Z(12) = \frac{1}{20}.$$

Η μέση τιμή της Z λαμβάνεται από τη ΣΜΠ της:

$$\begin{aligned}
\mathbf{E}[Z] &= \sum z p_Z(z) \\
&= 3\cdot\frac{1}{20} + 4\cdot\frac{1}{20} + 5\cdot\frac{2}{20} + 6\cdot\frac{2}{20} + 7\cdot\frac{4}{20} \\
&\quad + 8\cdot\frac{3}{20} + 9\cdot\frac{3}{20} + 10\cdot\frac{2}{20} + 11\cdot\frac{1}{20} + 12\cdot\frac{1}{20} \\
&= 7.55.
\end{aligned}$$

Εναλλακτικά, μπορούμε να υπολογίσουμε την $\mathbf{E}[Z]$ χρησιμοποιώντας τον τύπο

$$\mathbf{E}\big[Z\big] = \mathbf{E}\big[X\big] + 2\mathbf{E}\big[Y\big].$$

Από τις οριακές ΣΜΠ, οι οποίες δίνονται στο Σχ. 2.10, έχουμε

$$\mathbf{E}\big[X\big] = 1\cdot\frac{3}{20} + 2\cdot\frac{6}{20} + 3\cdot\frac{8}{20} + 4\cdot\frac{3}{20} = \frac{51}{20},$$

$$\mathbf{E}\big[Y\big] = 1\cdot\frac{3}{20} + 2\cdot\frac{7}{20} + 3\cdot\frac{7}{20} + 4\cdot\frac{3}{20} = \frac{50}{20},$$

άρα

$$\mathbf{E}[Z] = \frac{51}{20} + 2 \cdot \frac{50}{20} = 7.55.$$

Περισσότερες από Δύο Τυχαίες Μεταβλητές

Η από κοινού ΣΜΠ τριών τυχαίων μεταβλητών X, Y και Z ορίζεται ανάλογα με τα παραπάνω ως

$$p_{X,Y,Z}(x,y,z) = \mathrm{P}(X = x, Y = y, Z = z),$$

για όλες τις δυνατές τριάδες των αριθμητικών τιμών (x,y,z). Αντίστοιχες οριακές ΣΜΠ ορίζονται ανάλογα και υπολογίζονται μέσω σχέσεων όπως

$$p_{X,Y}(x,y) = \sum_{z} p_{X,Y,Z}(x,y,z),$$

και

$$p_X(x) = \sum_{y}\sum_{z} p_{X,Y,Z}(x,y,z).$$

Ο κανόνας της μέσης τιμής για συναρτήσεις δίνει

$$\mathbf{E}[g(X,Y,Z)] = \sum_{x}\sum_{y}\sum_{z} g(x,y,z)p_{X,Y,Z}(x,y,z),$$

και εάν η g είναι γραμμική και έχει τη μορφή $aX + bY + cZ + d$, τότε

$$\mathbf{E}[aX + bY + cZ + d] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c\mathbf{E}[Z] + d.$$

Επιπλέον, υπάρχουν προφανείς γενικεύσεις των ανωτέρω για περισσότερες από τρεις τυχαίες μεταβλητές. Για παράδειγμα, για οποιεσδήποτε τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ και οποιεσδήποτε σταθερές $a_1, a_2, \ldots, a_n$, έχουμε

$$\mathbf{E}[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1\mathbf{E}[X_1] + a_2\mathbf{E}[X_2] + \cdots + a_n\mathbf{E}[X_n].$$

Παράδειγμα 2.10. Η Μέση Τιμή της Διωνυμικής. Στο μάθημα των πιθανοτήτων έχουν εγγραφεί 300 φοιτητές και κάθε φοιτητής έχει πιθανότητα ίση με 1/3 να πάρει βαθμό μεγαλύτερο του 7, ανεξαρτήτως οποιουδήποτε άλλου φοιτητή. Ποια είναι η μέση τιμή της X, του συνολικού αριθμού των φοιτητών που παίρνουν βαθμό

μεγαλύτερο του 7; Έστω

$$X_i = \begin{cases} 1, & \text{εάν ο } i\text{-οστός φοιτητής παίρνει βαθμό μεγαλύτερο του 7,} \\ 0, & \text{αλλιώς.} \end{cases}$$

Άρα οι $X_1, X_2, \ldots, X_n$ είναι τυχαίες μεταβλητές Bernoulli με κοινή μέση τιμή $p = 1/3$. Το αθροισμά τους

$$X = X_1 + X_2 + \cdots + X_n$$

είναι ο αριθμός των φοιτητών που παίρνουν βαθμό μεγαλύτερο του 7. Επειδή η X είναι ο αριθμός των "επιτυχιών" σε n ανεξάρτητες επαναλήψεις, είναι μία διωνυμική τυχαία μεταβλητή με παραμέτρους n και p.

Χρησιμοποιώντας τη γραμμικότητα της X σαν συνάρτηση των X_i, έχουμε

$$\mathbf{E}[X] = \sum_{i=1}^{300} \mathbf{E}[X_i] = \sum_{i=1}^{300} \frac{1}{3} = 300 \cdot \frac{1}{3} = 100.$$

Εάν επαναλάβουμε τον υπολογισμό αυτό για ένα γενικό αριθμό φοιτητών n και πιθανότητα βαθμού μεγαλύτερου του 7 ίση με p, έχουμε

$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} p = np.$$

Παράδειγμα 2.11. Το Πρόβλημα των Καπέλων. Υποθέστε ότι n άτομα ρίχνουν τα καπέλα τους σε ένα κουτί και ακολούθως καθένας ανασύρει ένα καπέλο στην τύχη. (Κάθε καπέλο μπορεί να ανασυρθεί από ένα μόνο άτομο.) Ποια είναι η μέση τιμή της X, του αριθμού των ατόμων που ανασύρουν το δικό τους καπέλο;

Για το i-οστό άτομο, εισάγουμε την τυχαία μεταβλητή X_i η οποία έχει την τιμή 1 εάν το άτομο ανασύρει το δικό του καπέλο και έχει την τιμή 0 διαφορετικά. Επειδή $\mathrm{P}(X_i = 1) = 1/n$ και $\mathrm{P}(X_i = 0) = 1 - 1/n$, η μέση τιμή της X_i είναι

$$\mathbf{E}[X_i] = 1 \cdot \frac{1}{n} + 0 \cdot \left(1 - \frac{1}{n}\right) = \frac{1}{n}.$$

Έχουμε

$$X = X_1 + X_2 + \cdots + X_n,$$

έτσι ώστε

$$\mathbf{E}[X] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \cdots + \mathbf{E}[X_n] = n \cdot \frac{1}{n} = 1.$$

Σύνοψη Κύριων Σημείων για την από Κοινού ΣΜΠ

Έστω ότι οι X και Y είναι τυχαίες μεταβλητές που συνδέονται με το ίδιο πείραμα.

- Η **από κοινού ΣΜΠ** $p_{X,Y}$ των X και Y ορίζεται από την

$$p_{X,Y}(x,y) = \mathrm{P}(X = x, Y = y).$$

- Η **οριακή ΣΜΠ** των X και Y προκύπτει από την από κοινού ΣΜΠ, χρησιμοποιώντας τον τύπο

$$p_X(x) = \sum_y p_{X,Y}(x,y), \qquad p_Y(y) = \sum_x p_{X,Y}(x,y).$$

- Μία συνάρτηση $g(X,Y)$ των X και Y ορίζει μία νέα τυχαία μεταβλητή και

$$\mathbf{E}[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y).$$

  Εάν η g είναι γραμμική, της μορφής $aX + bY + c$, έχουμε

$$\mathbf{E}[aX + bY + c] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c.$$

- Τα παραπάνω έχουν φυσικές προεκτάσεις στην περίπτωση που έχουμε περισσότερες από δύο τυχαίες μεταβλητές.

2.6 ΔΕΣΜΕΥΣΗ

Σε αναλογία με τη συζήτηση στο Κεφάλαιο 1, δεσμευμένες πιθανότητες μπορούν να χρησιμοποιηθούν για να ερμηνεύσουμε την πληροφορία που διάφορα γεγονότα εμπεριέχουν σχετικά με τις τιμές μίας τυχαίας μεταβλητής. Σ' αυτή την παράγραφο εισάγουμε δεσμευμένες ΣΜΠ, δεδομένου κάποιου γεγονότος, ή δεδομένης της τιμής μίας τυχαίας μεταβλητής και συζητάμε τις ιδιότητές τους. Στην πραγματικότητα, δεν υπάρχουν πολλά καινούργια στοιχεία, μόνο κάποια λεπτομερής επεξεργασία εννοιών που είναι γνωστές από το Κεφάλαιο 1, καθώς και κάποιοι καινούργιοι συμβολισμοί.

Δεσμευμένη Τυχαία Μεταβλητή από ένα Γεγονός

Η **δεσμευμένη ΣΜΠ** μίας τυχαίας μεταβλητής X, δεδομένου ενός συγκεκριμένου γεγονότος A με $\mathrm{P}(A) > 0$, ορίζεται ως

$$p_{X|A}(x) = \mathrm{P}(X = x \,|\, A) = \frac{\mathrm{P}(\{X = x\} \cap A)}{\mathrm{P}(A)}.$$

Σημειώστε ότι τα γεγονότα $\{X = x\} \cap A$ είναι ξένα μεταξύ τους για διαφορετικές τιμές x, και η ένωσή τους είναι A, άρα

$$\mathrm{P}(A) = \sum_x \mathrm{P}(\{X = x\} \cap A).$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις, βλέπουμε ότι

$$\sum_x p_{X|A}(x) = 1,$$

άρα η $p_{X|A}$ είναι μία αποδεκτή ΣΜΠ.

Η δεσμευμένη ΣΜΠ υπολογίζεται με παρόμοιο τρόπο όπως οι αντίστοιχες μη δεσμευμένες ΣΜΠ: για να λάβουμε την $p_{X|A}(x)$, προσθέτουμε τις πιθανότητες των αποτελεσμάτων τα οποία αντιστοιχούν στην τιμή x **και** ανήκουν στο δεσμεύον γεγονός A και μετά κανονικοποιούμε διαιρώντας δια $\mathrm{P}(A)$.

Παράδειγμα 2.12. Έστω ότι η X είναι η τυχαία μεταβλητή της ρίψης ενός αμερόληπτου 6-εδρου ζαριού και έστω επίσης ότι A είναι το γεγονός ότι από τη ρίψη προκύπτει άρτιος αριθμός. Τότε εφαρμόζοντας τον προηγούμενο τύπο, έχουμε

$$\begin{aligned} p_{X|A}(k) &= \mathrm{P}(X = k \,|\, \mathrm{A}) \\ &= \frac{\mathrm{P}\{(X = k\} \cap \mathrm{A})}{\mathrm{P(A)}} \\ &= \begin{cases} 1/3, & \text{εάν } k = 2, 4, 6, \\ 0, & \text{αλλιώς.} \end{cases} \end{aligned}$$

Παράδειγμα 2.13. Ένας φοιτητής θα εξετασθεί σε ένα διαγώνισμα επανειλημμένα, μέχρι κάποιο μέγιστο αριθμό από n φορές, κάθε φορά με πιθανότητα p να το περάσει, ανεξαρτήτως του αριθμού των προηγουμένων προσπαθειών. Ποια είναι η ΣΜΠ του αριθμού των προσπαθειών δεδομένου ότι ο φοιτητής περνάει το διαγώνισμα;

Έστω A το γεγονός ότι ο φοιτητής πέρασε το διαγώνισμα (με το πολύ n προσπάθειες). Εισάγουμε την τυχαία μεταβλητή X, η οποία είναι ο αριθμός των προσπαθειών που χρειάζονται, εάν επιτρεπόταν ένας απεριόριστος αριθμός προσπαθειών. Τότε, η X

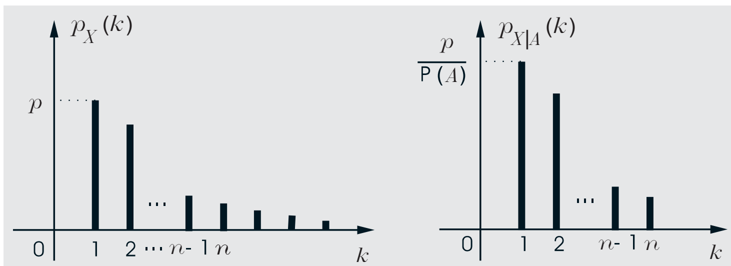
είναι μία γεωμετρική τυχαία μεταβλητή με παράμετρο p και $A = \{X \leq n\}$. Έχουμε

$$\mathrm{P}(A) = \sum_{m=1}^{n} (1-p)^{m-1} p,$$

και

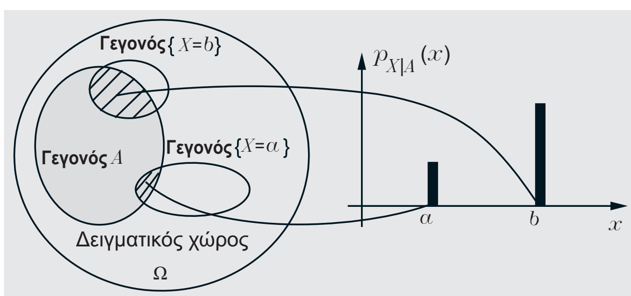
$$p_{X|A}(k) = \begin{cases} \dfrac{(1-p)^{k-1}p}{\sum_{m=1}^{n}(1-p)^{m-1}p}, & \text{εάν } k = 1, \ldots, n, \\ 0, & \text{αλλιώς,} \end{cases}$$

όπως απεικονίζεται στο Σχ. 2.11.



Σχήμα 2.11: Αναπαράσταση και υπολογισμός της δεσμευμένης ΣΜΠ $p_{X|A}(k)$ στο Παράδειγμα 2.13. Αρχίζουμε με την ΣΜΠ της X, μηδενίζουμε τις τιμές της ΣΜΠ για όλα τα k τα οποία δεν ανήκουν στο υπό δέσμευση γεγονός A και κανονικοποιούμε τις υπόλοιπες τιμές διαιρώντας διά $\mathrm{P}(A)$.

Το Σχήμα 2.12 δίνει μία πιο αφηρημένη αναπαράσταση της κατασκευής της δεσμευμένης ΣΜΠ.



Σχήμα 2.12: Αναπαράσταση και υπολογισμός της δεσμευμένης ΣΜΠ $p_{X|A}(x)$. Για κάθε x, προσθέτουμε τις πιθανότητες των ενδεχομένων της τομής $\{X = x\} \cap A$ και κανονικοποιούμε διαιρώντας δια $\mathrm{P}(A)$.

Δεσμευμένη Τυχαία Μεταβλητή από μία Άλλη Τυχαία Μεταβλητή

Έστω X και Y δύο τυχαίες μεταβλητές οι οποίες συνδέονται με το ίδιο πείραμα. Εάν γνωρίζουμε ότι η τιμή της Y είναι ίση με κάποιο συγκεκριμένο y [με $p_Y(y) > 0$], αυτό μας δίνει μερική πληροφορία για την τιμή της X. Η πληροφορία αυτή χρησιμοποιείται από τη **δεσμευμένη ΣΜΠ** $p_{X|Y}$ της X δεδομένης της Y, η οποία ορίζεται από την εξειδίκευση του ορισμού $p_{X|A}$ για γεγονότα A της μορφής $\{Y = y\}$:

$$p_{X|Y}(x \,|\, y) = \mathrm{P}(X = x \,|\, Y = y).$$

Εάν χρησιμοποιήσουμε τον ορισμό των δεσμευμένων πιθανοτήτων, έχουμε

$$p_{X|Y}(x \,|\, y) = \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

Ας κρατήσουμε σταθερή μία τιμή y με $p_Y(y) > 0$ και ας εξετάσουμε την $p_{X|Y}(x \,|\, y)$ ως συνάρτηση του x. Αυτή είναι μία αποδεκτή συνάρτηση ΣΜΠ της X: αναθέτει μη αρνητικές τιμές σε κάθε πιθανή τιμή x και αυτές οι τιμές αθροίζονται στο 1. Επιπλέον, η συνάρτηση αυτή του x έχει το ίδιο σχήμα όπως η $p_{X,Y}(x, y)$ εκτός του ότι είναι κανονικοποιημένη λόγω της διαίρεσής της δια της $p_Y(y)$, πράγμα το οποίο επιβάλλει την ιδιότητα της κανονικοποίησης

$$\sum_x p_{X|Y}(x \,|\, y) = 1.$$

Το Σχήμα 2.13 δίνει μία αναπαράσταση της δεσμευμένης ΣΜΠ.

Η δεσμευμένη ΣΜΠ συχνά υπολογίζεται από την από κοινού ΣΜΠ, χρησιμοποιώντας μία ακολουθιακή μέθοδο επίλυσης και τον τύπο

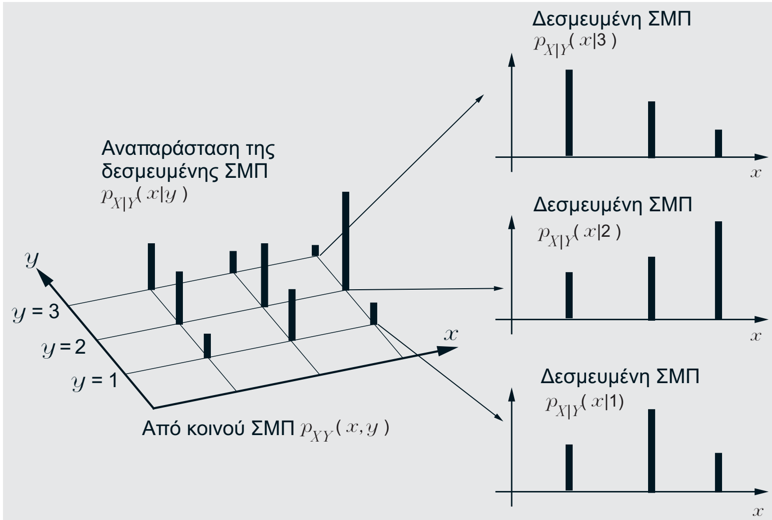
$$p_{X,Y}(x, y) = p_Y(y) p_{X|Y}(x \,|\, y),$$

ή τον τύπο

$$p_{X,Y}(x, y) = p_X(x) p_{Y|X}(y \,|\, x).$$

Η μέθοδος αυτή είναι εντελώς παρόμοια με τη χρήση του κανόνα του πολλαπλασιασμού του Κεφαλαίου 1.

Παράδειγμα 2.14. Μία καθηγήτρια απαντάει κάθε ερώτηση των φοιτητών της λανθασμένα με πιθανότητα 1/4, ανεξάρτητα από άλλες ερωτήσεις. Σε κάθε μάθημα η καθηγήτρια ερωτάται 0, 1, ή 2 ερωτήσεις με ίση πιθανότητα 1/3. Έστω X και Y ο αριθμός των ερωτήσεων που η καθηγήτρια ερωτάται και ο αριθμός των ερωτήσεων που απαντά λάθος, αντίστοιχα. Για να κατασκευάσουμε την από κοινού ΣΜΠ $p_{X,Y}(x, y)$, χρειάζεται να υπολογίσουμε όλες τις δυνατές πιθανότητες $\mathrm{P}(X = x, Y = y)$ για



Σχήμα 2.13: Αναπαράσταση μίας δεσμευμένης ΣΜΠ $p_{X|Y}(x \mid y)$. Για κάθε τιμή y, θεωρούμε την από κοινού ΣΜΠ κατά μήκος της γραμμής $Y = y$ και κανονικοποιούμε, έτσι ώστε

$$\sum_x p_{X|Y}(x \mid y) = 1.$$

όλους τους συνδυασμούς τιμών των x και y. Αυτό γίνεται με τη χρήση μίας ακολουθιακής περιγραφής του πειράματος και τον κανόνα του πολλαπλασιασμού, όπως φαίνεται στο Σχ. 2.14. Για παράδειγμα, για την περίπτωση όπου τίθεται ένα ερώτημα και η απάντησή του είναι λάθος, έχουμε

$$p_{X,Y}(1,1) = p_X(x)p_{Y|X}(y \mid x) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}.$$

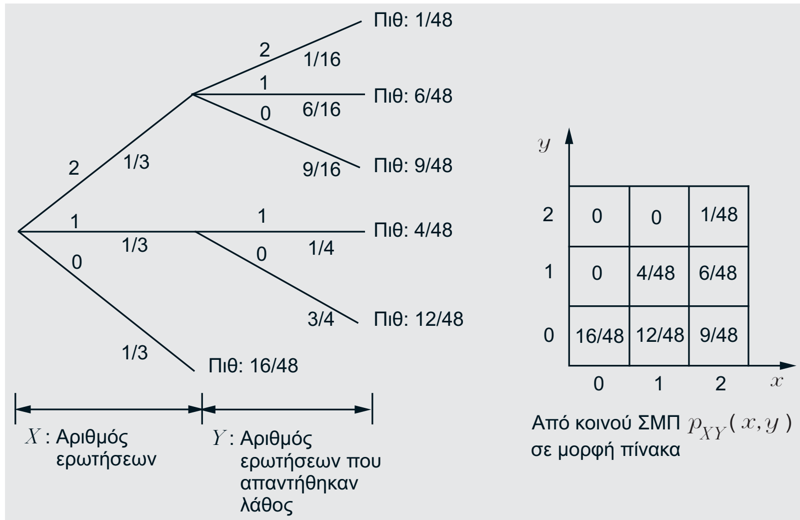
Η από κοινού ΣΜΠ μπορεί περιγραφεί με ένα διδιάστατο πίνακα, όπως φαίνεται στο Σχ. 2.14, και μπορεί να χρησιμοποιηθεί για τον υπολογισμό της πιθανότητας οποιουδήποτε γεγονότος μας ενδιαφέρει. Για παράδειγμα

$$\begin{aligned} \text{P}(\text{τουλάχιστον μία λάθος απάντηση}) &= p_{X,Y}(1,1) + p_{X,Y}(2,1) + p_{X,Y}(2,2) \\ &= \frac{4}{48} + \frac{6}{48} + \frac{1}{48}. \end{aligned}$$

.

Η δεσμευμένη ΣΜΠ μπορεί επίσης να χρησιμοποιηθεί για τον υπολογισμό των οριακών ΣΜΠ. Συγκεκριμένα χρησιμοποιώντας τους ορισμούς, έχουμε

$$p_X(x) = \sum_y p_{X,Y}(x,y) = \sum_y p_Y(y)p_{X|Y}(x \mid y).$$



Σχήμα 2.14: Υπολογισμός της από κοινού ΣΜΠ $p_{X,Y}(x, y)$ του Παραδείγματος 2.14.

Ο τύπος αυτός μας παρέχει μία μέθοδο "διαίρει και βασίλευε" για τον υπολογισμό των οριακών ΣΜΠ. Είναι βασικά ίδιος με το θεώρημα της συνολικής πιθανότητας που δίνεται στο Κεφάλαιο 1, αλλά παρουσιάζεται με διαφορετικό συμβολισμό. Το παρακάτω παράδειγμα δίνει μία επεξήγηση της μεθόδου.

Παράδειγμα 2.15. Θεωρήστε ένα πομπό ο οποίος στέλνει μηνύματα μέσω ενός δικτύου υπολογιστών. Ας ορίσουμε τις δύο παρακάτω τυχαίες μεταβλητές:

X : χρόνος αποστολής ενός δεδομένου μήνυματος,

Υ : το μήκος του δεδομένου μηνύματος.

Γνωρίζουμε τη ΣΜΠ του χρόνου αποστολής ενός μηνύματος το οποίο έχει δεδομένο μήκος και γνωρίζουμε τη ΣΜΠ του μήκους του μηνύματος. Θέλουμε να βρούμε τη (χωρίς δέσμευση) ΣΜΠ του χρόνου αποστολής του μηνύματος.

Υποθέτουμε ότι το μήκος ενός μηνύματος μπορεί να πάρει δύο δυνατές τιμές: $y = 10^2$ bytes με πιθανότητα 5/6 και $y = 10^4$ bytes με πιθανότητα 1/6, έτσι ώστε

$$p_Y(y) = \begin{cases} 5/6, & \text{εάν } y = 10^2, \\ 1/6, & \text{εάν } y = 10^4. \end{cases}$$

Υποθέτουμε ότι ο χρόνος αποστολής X του μηνύματος εξαρτάται από το μήκος Y και τη συμφόρηση του δικτύου την ώρα της αποστολής. Συγκεκριμένα ο χρόνος

αποστολής είναι $10^{-4}Y$ δευτερόλεπτα με πιθανότητα 1/2, $10^{-3}Y$ δευτερόλεπτα με πιθανότητα 1/3 και $10^{-2}Y$ δευτερόλεπτα με πιθανότητα 1/6. Άρα έχουμε

$$p_{X|Y}(x\,|\,10^2) = \begin{cases} 1/2, & \text{εάν } x = 10^{-2}, \\ 1/3, & \text{εάν } x = 10^{-1}, \\ 1/6, & \text{εάν } x = 1, \end{cases} \qquad p_{X|Y}(x\,|\,10^4) = \begin{cases} 1/2, & \text{εάν } x = 1, \\ 1/3, & \text{εάν } x = 10, \\ 1/6, & \text{εάν } x = 100. \end{cases}$$

Για να βρούμε τη ΣΜΠ της X χρησιμοποιούμε τον τύπο της συνολικής πιθανότητας

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x\,|\,y).$$

Έχουμε

$$p_X(10^{-2}) = \frac{5}{6}\cdot\frac{1}{2}, \qquad p_X(10^{-1}) = \frac{5}{6}\cdot\frac{1}{3}, \qquad p_X(1) = \frac{5}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{2},$$

$$p_X(10) = \frac{1}{6}\cdot\frac{1}{3}, \qquad p_X(100) = \frac{1}{6}\cdot\frac{1}{6}.$$

Τελικά, σημειώνουμε ότι μπορούμε να ορίσουμε ΣΜΠ που αφορούν περισσότερες από δύο τυχαίες μεταβλητές, όπως $p_{X,Y|Z}(x,y\,|\,z)$ ή $p_{X|Y,Z}\,(x\,|\,y,z)$. Οι έννοιες και οι μέθοδοι όπως τις έχουμε περιγράψει παραπάνω γενικεύονται εύκολα.

Σύνοψη Κύριων Σημείων για τη Δεσμευμένη ΣΜΠ

Έστω X και Y δύο τυχαίες μεταβλητές οι οποίες συνδέονται με το ίδιο πείραμα.

- Οι δεσμευμένες ΣΜΠ είναι παρόμοιες με τις συνήθεις ΣΜΠ, αλλά αναφέρονται σε ένα δειγματικό χώρο, όπου είναι γνωστό ότι το δεσμεύον γεγονός έχει συμβεί.

- Η δεσμευμένη ΣΜΠ της X δεδομένου του γεγονότος A με $\mathrm{P}(A) > 0$, ορίζεται ως

  $$p_{X|A}(x) = \mathrm{P}(X = x\,|\,A)$$

  και ικανοποιεί

  $$\sum_x p_{X|A}(x) = 1.$$

- Η δεσμευμένη ΣΜΠ της X δεδομένου $Y = y$ συνδέεται με την από

κοινού ΣΜΠ μέσω της σχέσης

$$p_{X,Y}(x,y) = p_Y(y)p_{X|Y}(x\,|\,y).$$

Ο τύπος αυτός είναι ανάλογος με τον κανόνα του πολλαπλασιασμού για τον υπολογισμό πιθανοτήτων και μπορεί να χρησιμοποιηθεί για να υπολογιστεί η από κοινού ΣΜΠ από τη δεσμευμένη ΣΜΠ.

- Η δεσμευμένη ΣΜΠ της X δεδομένης της Y μπορεί να χρησιμοποιηθεί για τον υπολογισμό της οριακής ΣΜΠ της X μέσω της σχέσης

$$p_X(x) = \sum_y p_Y(y)p_{X|Y}(x\,|\,y).$$

Η μέθοδος αυτή είναι ανάλογη με τη μέθοδο επίλυσης "διαίρει και βασίλευε" για τον υπολογισμό πιθανοτήτων χρησιμοποιώντας το θεώρημα της συνολικής πιθανότητας.

- Υπάρχουν φυσικές προεκτάσεις των παραπάνω που αφορούν περισσότερες από δύο τυχαίες μεταβλητές.

Δεσμευμένη Μέση Τιμή

Μία δεσμευμένη ΣΜΠ μπορεί να θεωρηθεί ως μία συνήθης ΣΜΠ ορισμένη σε ένα καινούργιο δειγματικό χώρο, ο οποίος προσδιορίζεται από το δεσμεύον γεγονός. Με τον ίδιο τρόπο, μία δεσμευμένη μέση τιμή είναι μία συνήθης μέση τιμή, εκτός του ότι αναφέρεται στο καινούργιο δειγματικό χώρο και όλες οι πιθανότητες και οι ΣΜΠ αντικαθιστώνται από αντίστοιχες δεσμευμένες. (Οι δεσμευμένες διασπορές μπορούν να θεωρηθούν με τον ίδιο τρόπο.) Παραθέτουμε παρακάτω τους κύριους ορισμούς και τα κύρια σημεία.

Σύνοψη Κύριων Σημείων για Δεσμευμένες Μέσες Τιμές

Έστω X και Y δύο τυχαίες μεταβλητές που αφορούν το ίδιο πείραμα.

- Η δεσμευμένη μέση τιμή της X δεδομένου του γεγονότος A με $\mathrm{P}(A) > 0$, ορίζεται ως

$$\mathbf{E}[X\,|\,A] = \sum_x x p_{X|A}(x).$$

Για μία συνάρτηση $g(X)$, έχουμε

$$\mathbf{E}[g(X) \mid A] = \sum_x g(x) p_{X|A}(x).$$

- Η δεσμευμένη μέση τιμή της X δεδομένης μίας τιμής y της Y ορίζεται ως

$$\mathbf{E}[X \mid Y = y] = \sum_x x p_{X|Y}(x \mid y).$$

- Έχουμε

$$\mathbf{E}[X] = \sum_y p_Y(y) \mathbf{E}[X \mid Y = y].$$

  Αυτό είναι το **θεώρημα της συνολικής μέσης τιμής**.

- Έστω $A_1, \ldots, A_n$ γεγονότα ξένα μεταξύ τους τα οποία αποτελούν μία διαμέριση του δειγματικού χώρου και υποθέτουμε ότι $\mathrm{P}(A_i) > 0$ για κάθε i. Τότε,

$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathrm{P}(A_i) \mathbf{E}[X \mid A_i].$$

- Έστω $A_1, \ldots, A_n$ γεγονότα ξένα μεταξύ τους τα οποία αποτελούν μία διαμέριση του γεγονότος B. Τότε, εφόσον $\mathrm{P}(A_i \cap B) > 0$ για κάθε i, έχουμε

$$\mathbf{E}[X \mid B] = \sum_{i=1}^{n} \mathrm{P}(A_i \mid B) \mathbf{E}[X \mid A_i \cap B].$$

Ας επαληθεύσουμε το θεώρημα της συνολικής μέσης τιμής, το οποίο βασικά λέει ότι "ο μη δεσμευμένος μέσος όρος προκύπτει από τον υπολογισμό του μέσου όρου των υπό δέσμευση μέσων όρων". Το θεώρημα αποδεικνύεται χρησιμοποιώντας τον τύπο της συνολικής πιθανότητας

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x \mid y),$$

και τον υπολογισμό

$$\begin{aligned}\mathbf{E}[X] &= \sum_x x p_X(x) \\ &= \sum_x x \sum_y p_Y(y) p_{X|Y}(x \mid y)\end{aligned}$$

$$= \sum_y p_Y(y) \sum_x x p_{X|Y}(x \mid y)$$
$$= \sum_y p_Y(y) \mathbf{E}[X \mid Y = y].$$

Η σχέση $\mathbf{E}[X] = \sum_{i=1}^n \mathrm{P}(A_i)\mathbf{E}[X \mid A_i]$ μπορεί να επαληθευτεί ως μία ειδική περίπτωση του θεωρήματος της συνολικής μέσης τιμής. Εισάγουμε την τυχαία μεταβλητή Y η οποία παίρνει την τιμή i εάν και μόνον εάν συμβαίνει το γεγονος A_i. Η ΣΜΠ της είναι

$$p_Y(i) = \begin{cases} \mathrm{P}(A_i), & \text{εάν } i = 1, 2, \ldots, n, \\ 0, & \text{αλλιώς.} \end{cases}$$

Το θεώρημα της συνολικής μέσης τιμής δίνει

$$\mathbf{E}[X] = \sum_{i=1}^n \mathrm{P}(A_i)\mathbf{E}[X \mid Y = i],$$

και εφόσον το γεγονός $\{Y = i\}$ είναι ακριβώς το A_i, έχουμε τη ζητούμενη έκφραση

$$\mathbf{E}[X] = \sum_{i=1}^n \mathrm{P}(A_i)\mathbf{E}[X \mid A_i].$$

Η τελευταία σχέση στον πίνακα σύνοψης είναι βασικά η ίδια με την προηγούμενη, αλλά εφαρμόζεται σε ένα δειγματικό χώρο όπου είναι γνωστό ότι το γεγονός B έχει συμβεί.

Το θεώρημα της συνολικής μέσης τιμής είναι ανάλογο με το θεώρημα της συνολικής πιθανότητας. Μπορεί να χρησιμοποιηθεί για τον υπολογισμό της μη δεσμευμένης μέσης τιμής $\mathbf{E}[X]$ από την δεσμευμένη ΣΜΠ ή μέση τιμή, χρησιμοποιώντας την ιδέα του "διαίρει και βασίλευε".

Παράδειγμα 2.16. Μηνύματα αποστέλλονται από έναν υπολογιστή στη Βοστώνη μέσω ενός δικτύου δεδομένων και κατευθύνονται προς τη Νέα Υόρκη με πιθανότητα 0.5, προς το Σικάγο με πιθανότητα 0.3 και προς το Σαν Φρανσίσκο με πιθανότητα 0.2. Ο χρόνος αποστολής X ενός μηνύματος είναι τυχαίος. Η μέση τιμή του είναι 0.05 δευτερόλεπτα εάν ο προορισμός του είναι η Νέα Υόρκη, 0.1 δευτερόλεπτα εάν ο προορισμός του είναι το Σικάγο και 0.3 δευτερόλεπτα εάν ο προορισμός του είναι το Σαν Φρανσίσκο. Τότε, η $\mathbf{E}[X]$ υπολογίζεται εύκολα χρησιμοποιώντας το θεώρημα της συνολικής μέσης τιμής

$$\mathbf{E}[X] = 0.5 \cdot 0.05 + 0.3 \cdot 0.1 + 0.2 \cdot 0.3 = 0.115 \text{ δευτερόλεπτα.}$$

Παράδειγμα 2.17. Μέση Τιμή και Διασπορά της Γεωμετρικής. Τρέχετε ένα πρόγραμμα λογισμικού επανειλημμένα και κάθε φορά υπάρχει πιθανότητα p να δουλέψει σωστά, ανεξάρτητα από προηγούμενες προσπάθειες. Ποια είναι η μέση τιμή και η διασπορά της X, του αριθμού των προσπαθειών μέχρι το πρόγραμμα να δουλέψει σωστά;

Αναγνωρίζουμε την X ως μία γεωμετρική τυχαία μεταβλητή με ΣΜΠ

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots.$$

Η μέση τιμή και η διασπορά της X είναι

$$\mathbf{E}[X] = \sum_{k=1}^{\infty} k(1-p)^{k-1}p, \qquad \mathrm{var}(X) = \sum_{k=1}^{\infty}(k - \mathbf{E}[X])^2(1-p)^{k-1}p,$$

αλλά ο υπολογισμός των άπειρων αθροισμάτων είναι δύσκολος. Εναλλακτικά, θα εφαρμόσουμε το θεώρημα της συνολικής μέσης τιμής, με $A_1 = \{X = 1\} = \{$η πρώτη προσπάθεια είναι επιτυχής$\}$, $A_2 = \{X > 1\} = \{$η πρώτη προσπάθεια είναι αποτυχημένη$\}$ και θα καταλήξουμε σε έναν πολύ πιο απλό υπολογισμό.

Εάν η πρώτη προσπάθεια είναι επιτυχής, έχουμε $X = 1$ και

$$\mathbf{E}[X \mid X = 1] = 1.$$

Εάν η πρώτη προσπάθεια αποτύχει ($X > 1$), έχουμε σπαταλήσει μία προσπάθεια και είμαστε πίσω εκεί πού αρχίσαμε. Άρα, ο μέσος αριθμός των προσπαθειών που απομένουν είναι $\mathbf{E}[X]$ και

$$\mathbf{E}[X \mid X > 1] = 1 + \mathbf{E}[X].$$

Συνεπώς,

$$\begin{aligned} \mathbf{E}[X] &= \mathrm{P}(X = 1)\mathbf{E}[X \mid X = 1] + \mathrm{P}(X > 1)\mathbf{E}[X \mid X > 1] \\ &= p + (1-p)\big(1 + \mathbf{E}[X]\big), \end{aligned}$$

από την οποία σχέση έχουμε

$$\mathbf{E}[X] = \frac{1}{p}.$$

Με παρόμοια λογική επίσης έχουμε

$$\mathbf{E}[X^2 \mid X = 1] = 1, \qquad \mathbf{E}[X^2 \mid X > 1] = \mathbf{E}\big[(1+X)^2\big] = 1 + 2\mathbf{E}[X] + \mathbf{E}[X^2],$$

έτσι ώστε

$$\mathbf{E}[X^2] = p \cdot 1 + (1-p)\big(1 + 2\mathbf{E}[X] + \mathbf{E}[X^2]\big),$$

από το οποίο λαμβάνουμε

$$\mathbf{E}[X^2] = \frac{1 + 2(1-p)\mathbf{E}[X]}{p},$$

και, χρησιμοποιώντας τον τύπο $\mathbf{E}[X] = 1/p$ που αποδείξαμε παραπάνω,

$$\mathbf{E}[X^2] = \frac{2}{p^2} - \frac{1}{p}.$$

Συμπεραίνουμε ότι

$$\text{var}(X) = \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Παράδειγμα 2.18. Το Παράδοξο των δύο Φακέλων. Αυτός είναι ένας πολυσυζητημένος γρίφος, ο οποίος σχετίζεται με ένα λεπτό μαθηματικό σημείο που αφορά δεσμευμένες μέσες τιμές.

Σας δίνονται δύο φάκελοι και σας λένε ότι ο ένας από αυτούς περιέχει m φορές περισσότερα χρήματα από τον άλλο, όπου m είναι ένας ακέραιος με $m > 1$. Ανοίγετε έναν από τους φακέλους και κοιτάζετε το περιεχόμενο. Μπορείτε να κρατήσετε το ποσό στο φάκελο, ή μπορείτε να τον ανταλλάξετε με το δεύτερο φάκελο και να κρατήσετε το ποσό που περιέχει. Ποια είναι η καλύτερη στρατηγική;

Ας παρουσιάσουμε ένα φαινομενικά λογικό επιχείρημα υπέρ της ανταλλαγής. Έστω ότι A είναι ο φάκελος που ανοίξατε και B ο φάκελος με τον οποίο μπορείτε να τον ανταλλάξετε. Έστω επίσης x και y τα ποσά μέσα στους A και B, αντίστοιχα. Τότε, ένα λογικοφανές επιχείρημα είναι ότι $y = x/m$ ή $y = mx$, με την ίδια πιθανότητα 1/2, άρα δεδομένου του x, η μέση τιμή του y είναι

$$\frac{1}{2} \cdot \frac{x}{m} + \frac{1}{2} \cdot mx = \frac{1}{2}\left(\frac{1}{m} + m\right)x = \frac{1+m^2}{2m}x > x,$$

καθώς $1 + m^2 > 2m$ για $m > 1$. Άρα, πάντα πρέπει να ανταλλάξετε τον A με τον B! Αλλά τότε, επειδή πρέπει να ανταλλάξετε τον A ανεξαρτήτως του ποσού που περιέχει, θα ήταν προτιμότερο να ανοίξετε κατ' αρχάς τον B, αλλά αν το κάνετε αυτό, θα πρέπει να τον ανταλλάξετε ξανά, κ.ο.κ..

Υπάρχουν δύο υποθέσεις, και οι δύο λανθασμένες σε κάποιο βαθμό, που ευθύνονται για αυτή την παράδοξη λογική.

(α) Δεν έχετε γνώση εκ των προτέρων του ποσού στους φακέλους, άρα δεδομένου του x, το μόνο πράγμα που ξέρετε για το y είναι ότι είναι $1/m$ ή m φορές x και δεν υπάρχει λόγος να υποθέσετε ότι το ένα είναι πιο πιθανό από το άλλο.

(β) Δεδομένων δύο τυχαίων μεταβλητών X και Y, οι οποίες εκφράζουν χρηματικά ποσά, εάν

$$\mathbf{E}[Y \mid X = x] > x,$$

για όλες τις δυνατές τιμές x της X, τότε η στρατηγική του να προτιμούμε πάντα το Y αποδίδει υψηλότερο μέσο χρηματικό κέρδος.

Ας εξετάσουμε πιο προσεκτικά τις υποθέσεις αυτές.

Η υπόθεση (α) δεν είναι σωστή, διότι βασίζεται σε ένα μοντέλο πιθανότητας το οποίο είναι ελλειπώς προσδιορισμένο. Πράγματι, σε οποιοδήποτε σωστό μοντέλο, όλα τα γεγονότα, συμπεριλαμβανομένων των δυνατών τιμών των X και Y, πρέπει να έχουν καλά ορισμένες πιθανότητες. Με τέτοιες γνώσεις πιθανότητας για την X και Y, η τιμή της X μπορεί να δώσει χρήσιμη πληροφορία για την Y. Για παράδειγμα, υποθέστε το παρακάτω μοντέλο πιθανότητας: κάποιος επιλέγει ένα ακέραιο αριθμό Z ευρώ από ένα γνωστό διάστημα $[\underline{z}, \overline{z}]$ σύμφωνα με κάποια κατανομή, τοποθετεί το ποσό αυτό σε ένα φάκελο που επέλεξε τυχαία και τοποθετεί m φορές το ίδιο ποσό στον άλλο φάκελο. Τότε ανοίγετε ένα από τους δύο φακέλους (με ίση πιθανότητα) και ελέγχετε το ποσό X που περιέχει. Εάν η X είναι μεγαλύτερη από το πάνω όριο $\overline{z}$, γνωρίζετε ότι η X είναι το μεγαλύτερο των δύο ποσών και, επομένως, δεν χρειάζεται να αλλάξετε. Αφετέρου, για κάποιες άλλες τιμές της X, όπως το κάτω όριο $\underline{z}$, θα πρέπει να ανταλλάξετε τους φακέλους. Επομένως, σ' αυτό το μοντέλο, η επιλογή της ανταλλαγής εξαρτάται από την τιμή της X. Χωρίς να ακριβολογούμε, εάν γνωρίζετε τις πιθανότητες των τιμών της X, μπορείτε να κρίνετε αν το ποσό που βρέθηκε στον Α είναι σχετικά μικρό ή σχετικά μεγάλο και ανάλογα να ανταλλάξετε τους φακέλους.

Από μαθηματικής άποψης, σε ένα σωστό μοντέλο πιθανοτήτων, πρέπει να έχουμε την από κοινού ΣΜΠ για τις τυχαίες μεταβλητές X και Y, τα ποσά των φακέλων Α και Β, αντίστοιχα. Η από κοινού αυτή ΣΜΠ προσδιορίζεται με την εισαγωγή μίας ΣΜΠ p_Z για την τυχαία μεταβλητή Z, το ελάχιστο του ποσού των δύο φακέλων. Τότε, για όλα τα z,

$$p_{X,Y}(mz, z) = p_{X,Y}(z, mz) = \frac{1}{2} p_Z(z),$$

και

$$p_{X,Y}(x, y) = 0,$$

για κάθε (x, y) το οποίο δεν είναι της μορφής (mz, z) ή (z, mz). Με αυτό τον προσδιορισμό της $p_{X,Y}(x, y)$ και δεδομένου ότι $X = x$, κάποιος μπορεί να χρησιμοποιήσει τον κανόνα

$$\text{ανταλλάξτε εάν και μόνο εάν } \mathbf{E}[Y \mid X = x] > x.$$

Σύμφωνα με αυτό τον κανόνα, κάποιος μπορεί να ανταλλάξει ή όχι τους φακέλους, ανάλογα με την τιμή της X, όπως έχουμε δείξει νωρίτερα.

Αληθεύει ότι με το μοντέλο πιθανότητας όπως έχει περιγραφεί παραπάνω καθώς και τον προηγούμενο κανόνα λήψης απόφασης, θα πρέπει να κάνετε ανταλλαγή για μερικές τιμές x, αλλά όχι για άλλες; Κανονικά ναι, όπως έχει επεξηγηθεί στο προηγούμενο παράδειγμα όπου η Z παίρνει τιμές σε ένα φραγμένο διάστημα. Παρόλα αυτά, υπάρχει μία παράδοξη περίπτωση όπου εξαιτίας μίας παθολογικής μαθηματικής συγκυρίας, πάντα θα κάνετε ανταλλαγή!

Ένα αμερόληπτο νόμισμα ρίχνεται μέχρι να έρθει κορώνα. Έστω N ο αριθμός των ρίψεων. Τότε, m^N ευρώ τοποθετούνται σε ένα φάκελο και m^{N-1} ευρώ τοποθετούνται σε ένα άλλο. Έστω X το ποσό στο φάκελο που ανοίξατε (φάκελο Α) και έστω Y το ποσό στον άλλο φάκελο (φάκελο Β).

Τώρα, εάν ο Α περιέχει 1 ευρώ, ξεκάθαρα ο Β περιέχει m ευρώ, άρα πρέπει να

ανταλλάξετε τους φακέλους. Εάν, από την άλλη μεριά, ο Α περιέχει m^n ευρώ, όπου $n > 0$, τότε ο Β περιέχει είτε m^{n-1} ή m^{n+1} ευρώ. Καθώς το N έχει μία γεωμετρική ΣΜΠ, έχουμε

$$\frac{\mathrm{P}(Y = m^{n+1} \,|\, X = m^n)}{\mathrm{P}(Y = m^{n-1} \,|\, X = m^n)} = \frac{\mathrm{P}(Y = m^{n+1},\, X = m^n)}{\mathrm{P}(Y = m^{n-1},\, X = m^n)} = \frac{\mathrm{P}(N = n+1)}{\mathrm{P}(N = n)} = \frac{1}{2}.$$

Άρα

$$\mathrm{P}(Y = m^{n-1} \,|\, X = m^n) = \frac{2}{3}, \qquad \mathrm{P}(Y = m^{n+1} \,|\, X = m^n) = \frac{1}{3},$$

και

$$\mathbf{E}[\text{ποσό στο Β} \,|\, X = m^n] = \frac{2}{3} \cdot m^{n-1} + \frac{1}{3} \cdot m^{n+1} = \frac{2 + m^2}{3m} \cdot m^n.$$

Έχουμε $(2 + m^2)/3m > 1$ εάν και μόνο εάν $m^2 - 3m + 2 > 0$ ή $(m-1)(m-2) > 0$. Άρα εάν $m > 2$, τότε

$$\mathbf{E}[\text{ποσό στο Β} \,|\, X = m^n] > m^n,$$

και για να μεγιστοποιήστε το αναμενόμενο χρηματικό κέρδος θα πρέπει πάντα να ανταλλάξετε με το Β!

Οδηγούμαστε σ' αυτό το παράδοξο επειδή

$$\mathbf{E}[Y \,|\, X = x] > x, \qquad \text{για όλα τα } x,$$

που φαινομενικά, με βάση το θεώρημα της συνολικής μέσης τιμής, οδηγεί στο συμπέρασμα ότι $\mathbf{E}[Y] > \mathbf{E}[X]$. Ωστόσο, αυτό δεν μπορεί να αληθεύει, διότι η X και η Y έχουν ίδιες ΣΜΠ. Αυτό που συμβαίνει στην προκείμενη περίπτωση είναι ότι

$$\mathbf{E}[Y] = \mathbf{E}[X] = \infty,$$

το οποίο δεν αντικρούει κατ' ανάγκη τη σχέση $\mathbf{E}[Y \,|\, X = x] > x$ για όλα τα x.

Το συμπέρασμα είναι ότι ο κανόνας λήψης απόφασης σύμφωνα με τον οποίο γίνεται ανταλλαγή εάν και μόνο εάν $\mathbf{E}[Y \,|\, X = x] > x$ δεν βελτιώνει το αναμενόμενο χρηματικό κέρδος στην περίπτωση όπου $\mathbf{E}[Y] = \mathbf{E}[X] = \infty$ και το φαινομενικό παράδοξο εξηγείται.

2.7 ΑΝΕΞΑΡΤΗΣΙΑ

Θα συζητήσουμε τώρα έννοιες ανεξαρτησίας τυχαίων μεταβλητών. Οι έννοιες αυτές είναι ανάλογες με τις έννοιες ανεξαρτησίας μεταξύ δύο γεγονότων (βλ. Κεφάλαιο 1). Αναπτύσσονται με την εισαγωγή κατάλληλων γεγονότων που α-

φορούν τις δυνατές τιμές διαφόρων τυχαίων μεταβλητών και με τη θεώρηση της ανεξαρτησίας των γεγονότων αυτών.

Ανεξαρτησία μίας Τυχαίας Μεταβλητής από ένα Γεγονός

Η ανεξαρτησία μίας τυχαίας μεταβλητής από ένα γεγονός είναι παρόμοια με την ανεξαρτησία δύο γεγονότων. Η ιδέα είναι ότι η γνώση ότι το δεσμεύον γεγονός έχει συμβεί δεν μας δίνει καμία καινούργια πληροφορία για την τιμή της τυχαίας μεταβλητής. Ακριβέστερα, λέμε ότι η τυχαία μεταβλητή X είναι **ανεξάρτητη από το γεγονός** A εάν

$$\mathrm{P}(X = x \text{ και } A) = \mathrm{P}(X = x)\mathrm{P}(A) = p_X(x)\mathrm{P}(A), \qquad \text{για όλες τις τιμές } x,$$

το οποίο είναι το ίδιο με το να απαιτούμε ότι τα δύο γεγονότα $\{X = x\}$ και A είναι ανεξάρτητα, για οποιαδήποτε τιμή x. Από τον ορισμό της δεσμευμένης ΣΜΠ, έχουμε

$$\mathrm{P}(X = x \text{ και } A) = p_{X|A}(x)\mathrm{P}(A),$$

έτσι ώστε, εφόσον $\mathrm{P}(A) > 0$, η ανεξαρτησία είναι ισοδύναμη με τη συνθήκη

$$p_{X|A}(x) = p_X(x), \qquad \text{για όλες τις τιμές } x.$$

Παράδειγμα 2.19. Θεωρήστε δύο ανεξάρτητες ρίψεις ενός αμερόληπτου νομίσματος. Έστω X ο αριθμός των κορωνών και έστω A το γεγονός ότι ο αριθμός των κορωνών είναι άρτιος. Η (χωρίς δέσμευση) ΣΜΠ της X είναι

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

και $\mathrm{P}(A) = 1/2$. Η δεσμευμένη ΣΜΠ προκύπτει από τον ορισμό $p_{X|A}(x) = \mathrm{P}\big(X = x \text{ και } A\big)/\mathrm{P}(A)$:

$$p_{X|A}(x) = \begin{cases} 1/2, & \text{εάν } x = 0, \\ 0, & \text{εάν } x = 1, \\ 1/2, & \text{εάν } x = 2. \end{cases}$$

Προφανώς, η X δεν είναι ανεξάρτητη από το A, εφόσον οι ΣΜΠ p_X και $p_{X|A}$ είναι διαφορετικές. Για ένα παράδειγμα μίας τυχαίας μεταβλητής που είναι ανεξάρτητη του A, θεωρήστε την τυχαία μεταβλητή που παίρνει την τιμή 0 εάν η πρώτη ρίψη είναι κορώνα και την τιμή 1 εάν η πρώτη ρίψη είναι γράμματα. Αυτό είναι διαισθητικά ξεκάθαρο και μπορεί να επαληθευτεί χρησιμοποιώντας τον ορισμό της ανεξαρτησίας.

Ανεξαρτησία Τυχαίων Μεταβλητών

Η έννοια της ανεξαρτησίας δύο τυχαίων μεταβλητών είναι παρόμοια με την ανεξαρτησία μίας τυχαίας μεταβλητής από ένα γεγονός. Λέμε ότι δύο **τυχαίες μεταβλητές** X και Y είναι **ανεξάρτητες** εάν

$$p_{X,Y}(x,y) = p_X(x)\, p_Y(y), \qquad \text{για όλα τα } x, y.$$

Η σχέση αυτή είναι ισοδύναμη με την απαίτηση ότι τα δύο γεγονότα $\{X = x\}$ και $\{Y = y\}$ είναι ανεξάρτητα για κάθε x και y. Τέλος, ο τύπος $p_{X,Y}(x,y) = p_{X|Y}(x \,|\, y)p_Y(y)$ δείχνει ότι η ανεξαρτησία είναι ισοδύναμη με τη συνθήκη

$$p_{X|Y}(x \,|\, y) = p_X(x), \qquad \text{για κάθε } y \text{ με } p_Y(y) > 0 \text{ και όλα τα } x.$$

Διαισθητικά, ανεξαρτησία σημαίνει ότι η τιμή της Y δεν παρέχει πληροφορία για την τιμή της X.

Παρομοίως μπορούμε να θεωρήσουμε τη δεσμευμένη ανεξαρτησία δύο τυχαίων μεταβλητών, δεδομένου ενός γεγονότος A με $\mathrm{P}(A) > 0$. Το γεγονός A της δέσμευσης ορίζει ένα καινούργιο δειγματικό χώρο και όλες οι πιθανότητες (ή οι ΣΜΠ) πρέπει να αντικατασταθούν από τις αντίστοιχες δεσμευμένες. Για παράδειγμα, οι X και Y λέγονται **ανεξάρτητες υπό δέσμευση** εάν δεδομένου ενός γεγονότος A με θετική πιθανότητα, ισχύει

$$\mathrm{P}(X = x, Y = y \,|\, A) = \mathrm{P}(X = x \,|\, A)\mathrm{P}(Y = y \,|\, A), \quad \text{για όλα τα } x \text{ και } y,$$

ή, σύμφωνα με το συμβολισμό του κεφαλαίου αυτού,

$$p_{X,Y|A}(x,y) = p_{X|A}(x)p_{Y|A}(y), \qquad \text{για όλες τις τιμές } x \text{ και } y.$$

Ξανά, αυτό είναι ισοδύναμο με το

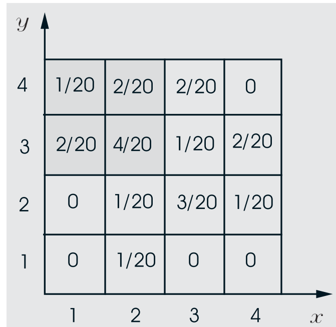
$$p_{X|Y,A}(x \,|\, y) = p_{X|A}(x) \qquad \text{για όλα τα } x \text{ και } y \text{ για τα οποία } p_{Y|A}(y) > 0.$$

Όπως στην περίπτωση των γεγονότων (Παράγραφος 1.5), η δεσμευμένη ανεξαρτησία δεν συνεπάγεται ανεξαρτησία χωρίς δέσμευση και αντίστροφα. Αυτό επεξηγείται στο παράδειγμα του Σχ. 2.15.

Εάν οι X και Y είναι ανεξάρτητες τυχαίες μεταβλητές, τότε

$$\mathbf{E}[XY] = \mathbf{E}[X]\,\mathbf{E}[Y],$$

όπως αποδεικνύεται από τον ακόλουθο υπολογισμό:



Σχήμα 2.15: Παράδειγμα που επεξηγεί ότι η δεσμευμένη ανεξαρτησία μπορεί να μην συνεπάγεται ανεξαρτησία χωρίς δέσμευση. Για την από κοινού ΣΜΠ που φαίνεται στο σχήμα, οι τυχαίες μεταβλητές X και Y δεν είναι ανεξάρτητες. Για παράδειγμα, έχουμε

$$p_{X|Y}(1\,|\,1) = \mathrm{P}(X = 1\,|\,Y = 1) = 0 \neq \mathrm{P}(X = 1) = p_X(1).$$

Αφετέρου, η δέσμευση από το γεγονός $A = \{X \leq 2,\ Y \geq 3\}$ (το σκιαγραφημένο σύνολο στο σχήμα) δείχνει ότι οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες. Συγκεκριμένα, έχουμε

$$p_{X|Y,A}(x\,|\,y) = \begin{cases} 1/3, & \text{εάν } x = 1, \\ 2/3, & \text{εάν } x = 2, \end{cases}$$

και για τις δύο τιμές $y = 3$ και $y = 4$.

$$\begin{aligned}
\mathbf{E}[XY] &= \sum_x \sum_y xy p_{X,Y}(x, y) \\
&= \sum_x \sum_y xy p_X(x) p_Y(y) \qquad\qquad \text{(λόγω ανεξαρτησίας)} \\
&= \sum_x x p_X(x) \sum_y y p_Y(y) \\
&= \mathbf{E}[X]\,\mathbf{E}[Y].
\end{aligned}$$

Παρόμοιος υπολογισμός αποδεικνύει επίσης ότι εάν οι X και Y είναι ανεξάρτητες, τότε

$$\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)]\mathbf{E}[h(Y)],$$

για οποιεσδήποτε συναρτήσεις g και h. Πράγματι, αυτό ακολουθεί άμεσα, εφόσον αντιληφθούμε ότι εάν οι X και Y είναι ανεξάρτητες, τότε το ίδιο ισχύει και για τις $g(X)$ και $h(Y)$. Αυτό είναι διαισθητικά σαφές και η μαθηματική επαλήθευσή του αφήνεται σαν πρόβλημα στο τέλος του κεφαλαίου.

Θεωρήστε τώρα το άθροισμα $X + Y$ δύο ανεξάρτητων τυχαίων μεταβλητών X και Y και ας υπολογίσουμε τη διασπορά του. Εφόσον η διασπορά μίας τυχαίας

μεταβλητής δεν αλλάζει όταν στην τυχαία μεταβλητή προστεθεί μία σταθερά, θεωρούμε τις τυχαίες μεταβλητές, $\tilde{X} = X - \mathbf{E}[X]$ και $\tilde{Y} = Y - \mathbf{E}[Y]$ που έχουν μηδενική μέση τιμή. Έχουμε τότε

$$\begin{aligned}
\text{var}(X+Y) &= \text{var}(\tilde{X}+\tilde{Y}) \\
&= \mathbf{E}[(\tilde{X}+\tilde{Y})^2] \\
&= \mathbf{E}[\tilde{X}^2 + 2\tilde{X}\,\tilde{Y} + \tilde{Y}^2] \\
&= \mathbf{E}[\tilde{X}^2] + 2\mathbf{E}[\tilde{X}\,\tilde{Y}] + \mathbf{E}[\tilde{Y}^2] \\
&= \mathbf{E}[\tilde{X}^2] + \mathbf{E}[\tilde{Y}^2] \\
&= \text{var}(\tilde{X}) + \text{var}(\tilde{Y}) \\
&= \text{var}(X) + \text{var}(Y).
\end{aligned}$$

Στον προηγούμενο υπολογισμό χρησιμοποιήσαμε την ιδιότητα $\mathbf{E}[\tilde{X}\,\tilde{Y}] = 0$, η οποία δικαιολογείται ως εξής. Οι τυχαίες μεταβλητές $\tilde{X} = X - \mathbf{E}[X]$ και $\tilde{Y} = Y - \mathbf{E}[Y]$ είναι ανεξάρτητες (διότι είναι συναρτήσεις ανεξαρτήτων τυχαίων μεταβλητών X και Y) και εφόσον έχουν μηδενικές μέσες τιμές, έχουμε

$$\mathbf{E}[\tilde{X}\,\tilde{Y}] = \mathbf{E}[\tilde{X}]\,\mathbf{E}[\tilde{Y}] = 0.$$

Το συμπέρασμα είναι ότι η διασπορά του αθροίσματος δύο **ανεξαρτήτων** τυχαίων μεταβλητών ισούται με το άθροισμα των διασπορών τους. Σε αντίθεση, σημειώστε ότι η μέση τιμή του αθροίσματος δύο τυχαίων μεταβλητών είναι πάντα ίση με το άθροισμα των μέσων τιμών τους, ακόμα και αν δεν είναι ανεξάρτητες.

Σύνοψη Κύριων Σημείων για Ανεξάρτητες Τυχαίες Μεταβλητές

Έστω ότι A είναι ένα γεγονός με $\mathrm{P}(A) > 0$ και έστω X και Y τυχαίες μεταβλητές σχετικές με το ίδιο πείραμα.

- Η X είναι ανεξάρτητη από το γεγονός A εάν

  $$p_{X|A}(x) = p_X(x), \qquad \text{για κάθε } x,$$

  δηλαδή, εάν για όλα τα x, τα γεγονότα $\{X = x\}$ και A είναι ανεξάρτητα.

- Οι X και Y είναι ανεξάρτητες εάν για όλα τα δυνατά ζεύγη (x, y), τα γεγονότα $\{X = x\}$ και $\{Y = y\}$ είναι ανεξάρτητα, ή ισοδύναμα

  $$p_{X,Y}(x, y) = p_X(x)p_Y(y), \qquad \text{για κάθε } x, y.$$

• Εάν οι X και Y είναι ανεξάρτητες τυχαίες μεταβλητές, τότε

$$\mathbf{E}[XY] = \mathbf{E}[X]\,\mathbf{E}[Y].$$

Επιπλέον, για οποιεσδήποτε συναρτήσεις g και h, οι τυχαίες μεταβλητές $g(X)$ και $h(Y)$ είναι ανεξάρτητες και έχουμε

$$\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)]\,\mathbf{E}[h(Y)].$$

• Εάν οι X και Y είναι ανεξάρτητες, τότε

$$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y).$$

Ανεξαρτησία Πολλών Τυχαίων Μεταβλητών

Η προηγούμενη συζήτηση επεκτείνεται στην περίπτωση των περισσοτέρων των δύο τυχαίων μεταβλητών. Για παράδειγμα, τρεις τυχαίες μεταβλητές X, Y και Z λέγονται ανεξάρτητες εάν

$$p_{X,Y,Z}(x,y,z) = p_X(x)p_Y(y)p_Z(z), \qquad \text{για κάθε } x, y, z.$$

Εάν οι X, Y και Z είναι ανεξάρτητες τυχαίες μεταβλητές, τότε οποιεσδήποτε τρεις τυχαίες μεταβλητές της μορφής $f(X)$, $g(Y)$ και $h(Z)$, είναι επίσης ανεξάρτητες. Παρόμοια, οποιεσδήποτε δύο τυχαίες μεταβλητές της μορφής $g(X,Y)$ και $h(Z)$ είναι ανεξάρτητες. Όμως δύο τυχαίες μεταβλητές της μορφής $g(X,Y)$ και $h(Y,Z)$ συνήθως δεν είναι ανεξάρτητες, επειδή και οι δύο επηρεάζονται από την Y. Ιδιότητες σαν τις παραπάνω είναι συνήθως διαισθητικά ξεκάθαρες. Είναι δυνατόν να επαληθευτούν τυπικά, αλλά αυτό είναι μερικές φορές επίπονο. Ευτυχώς, υπάρχει γενική συμφωνία μεταξύ διαίσθησης και μαθηματκής ορθότητας. Αυτό αποτελεί μία ένδειξη ότι οι ορισμοί μας σχετικά με την ανεξαρτησία αντικατοπτρίζουν ικανοποιητικά τις ερμηνείες που δίνουμε.

Διασπορά του Αθροίσματος Ανεξάρτητων Τυχαίων Μεταβλητών

Αθροίσματα ανεξάρτητων τυχαίων μεταβλητών είναι ιδιαίτερα σημαντικά σε πολλές περιπτώσεις. Για παράδειγμα, προκύπτουν σε στατιστικές εφαρμογές όπου παίρνουμε το “μέσο όρο” από ανεξάρτητες μετρήσεις με σκοπό την ελαχιστοποίηση των επιδράσεων των λαθών των μετρήσεων. Επίσης προκύπτουν όταν έχουμε να κάνουμε με αθροιστικές επιδράσεις διαφόρων ανεξάρτητων πηγών τυ-

χαιότητας. Παραθέτουμε μερικά παραδείγματα και θα επιστρέψουμε ξανά στο θέμα αυτό στα κεφάλαια που ακολουθούν.

Στο παρακάτω παράδειγμα, θα κάνουμε χρήση της ακόλουθης ιδιότητας κλειδί. Εάν οι $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές, τότε

$$\text{var}(X_1 + X_2 + \cdots + X_n) = \text{var}(X_1) + \text{var}(X_2) + \cdots + \text{var}(X_n).$$

Αυτό μπορεί να επαληθευτεί με επανειλημμένη χρήση του τύπου $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$ για δύο ανεξάρτητες τυχαίες μεταβλητές X και Y.

Παράδειγμα 2.20. Διασπορά της Διωνυμικής. Θεωρούμε n ανεξάρτητες ρίψεις ενός νομίσματος, με κάθε ρίψη να έχει πιθανότητα p να προκύψει κορώνα. Για κάθε i, έστω X_i μία τυχαία μεταβλητή Bernoulli ίση με 1 εάν η i-οστή ρίψη προκύπτει κορώνα και 0 διαφορετικά. Τότε, η $X = X_1 + X_2 + \cdots + X_n$ είναι μία διωνυμική τυχαία μεταβλητή. Λόγω της ανεξαρτησίας των ρίψεων του νομίσματος, οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ είναι ανεξάρτητες, άρα

$$\text{var}(X) = \sum_{i=1}^{n} \text{var}(X_i) = np(1-p).$$

Οι τύποι της μέσης τιμής και της διασποράς ενός γραμμικού αθροίσματος τυχαίων μεταβλητών είναι η βάση πολλών στατιστικών διαδικασιών οι οποίες προσεγγίζουν τη μέση τιμή μίας τυχαίας μεταβλητής βάσει του “μέσου όρου” πολλών ανεξάρτητων δειγμάτων. Το παρακάτω παράδειγμα περιγράφει μία τυπική περίπτωση.

Παράδειγμα 2.21. Μέση τιμή και Διασπορά της Δειγματικής Μέσης Τιμής. Θέλουμε να εκτιμήσουμε την πρόθεση ψήφου υπέρ ενός πολιτικού που τον ονομάζουμε Β. Για το σκοπό αυτό, ρωτάμε n άτομα τα οποία επιλέγονται τυχαία από τους ψηφοφόρους και έστω X_i η τυχαία μεταβλητή που αντιπροσωπεύει την απάντηση του i-οστού ατόμου:

$$X_i = \begin{cases} 1, & \text{εάν το } i\text{-οστό είναι υπέρ του Β}, \\ 0, & \text{εάν το } i\text{-οστό άτομο δεν είναι υπέρ του Β}. \end{cases}$$

Μοντελοποιούμε τις $X_1, X_2, \ldots, X_n$ ως ανεξάρτητες τυχαίες μεταβλητές Bernoulli με κοινή μέση τιμή p και διασπορά $p(1-p)$. Όπως είναι φυσικό, θεωρούμε ότι το p είναι η πραγματική πρόθεση ψήφου υπέρ του Β. Παίρνουμε το “μέσο όρο” των απαντήσεων, δηλαδή υπολογίζουμε τη **δειγματική μέση τιμή**, S_n, η οποία ορίζεται ως

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Η τυχαία μεταβλητή S_n είναι η πρόθεση ψήφου υπέρ του B στο δείγμα των n ατόμων.

Χρησιμοποιώντας τη γραμμικότητα της S_n ως συνάρτηση των X_i, έχουμε

$$\mathbf{E}[S_n] = \sum_{i=1}^{n} \frac{1}{n}\mathbf{E}[X_i] = \frac{1}{n}\sum_{i=1}^{n} p = p,$$

και λόγω της ανεξαρτησίας των $X_1, \ldots, X_n$,

$$\text{var}(S_n) = \sum_{i=1}^{n} \frac{1}{n^2}\text{var}(X_i) = \frac{p(1-p)}{n}.$$

Η δειγματική μέση τιμή S_n μπορεί να θεωρηθεί "καλή" εκτίμηση της πρόθεσης ψήφου διότι έχει τη σωστή μέση τιμή p και η ακρίβειά της, όπως φαίνεται από τη διασπορά, βελτιώνεται καθώς το μέγεθος του δείγματος n αυξάνει.

Παρατηρήστε ότι και αν ακόμα οι τυχαίες μεταβλητές X_i δεν είναι Bernoulli, ο ίδιος υπολογισμός δίνει

$$\text{var}(S_n) = \frac{\text{var}(X)}{n},$$

εφόσον οι X_i είναι ανεξάρτητες, με κοινή μέση τιμή $\mathbf{E}[X]$ και διασπορά $\text{var}(X)$. Άρα, ξανά, η δειγματική μέση τιμή αποτελεί μία καλή εκτίμηση (από την άποψη της διασποράς) της πραγματικής μέσης τιμής $\mathbf{E}[X]$, όταν το μέγεθος του δείγματος n αυξάνει. Θα επισκεφτούμε ξανά τις ιδιότητες της δειγματικής μέσης τιμής και θα τις συζητήσουμε με περισσότερη λεπτομέρεια στο Κεφάλαιο 7, σε σχέση με τους νόμους των μεγάλων αριθμών.

Παράδειγμα 2.22. Εκτίμηση Πιθανότητας μέσω Προσομοίωσης. Συχνά στην πράξη, οι αναλυτικοί υπολογισμοί των πιθανοτήτων κάποιου γεγονότος είναι πολύ δύσκολοι. Ωστόσο, εάν έχουμε ένα φυσικό ή υπολογιστικό μοντέλο που μπορεί να παράγει αποτελέσματα ενός δεδομένου πειράματος σύμφωνα με τις πραγματικές τους πιθανότητες, μπορούμε να χρησιμοποιήσουμε προσομοίωση για να υπολογίσουμε με μεγάλη ακρίβεια την πιθανότητα οποιουδήποτε δεδομένου γεγονότος A. Συγκεκριμένα, δημιουργούμε με το μοντέλο μας n ανεξάρτητα αποτελέσματα, καταγράφουμε τον αριθμό m των αποτελεσμάτων που ανήκουν στο γεγονός A που μας ενδιαφέρει και προσεγγίζουμε την $\text{P}(A)$ με m/n. Για παράδειγμα, για να υπολογίσουμε την πιθανότητα $p = \text{P}(\text{κορώνα})$ ενός νομίσματος, ρίχνουμε το νόμισμα n φορές και προσεγγίζουμε το p με το πηλίκο (αριθμός κορωνών που καταγράψαμε)$/n$.

Για να δούμε πόσο ακριβής είναι η διαδικασία αυτή, θεωρούμε n ανεξάρτητες τυχαίες μεταβλητές Bernoulli $X_1, \ldots, X_n$, κάθε μία με

$$p_{X_i}(k) = \begin{cases} \text{P}(A), & \text{εάν } k = 1, \\ 1 - \text{P}(A), & \text{εάν } k = 0. \end{cases}$$

Στο πλαίσιο της προσομοίωσης, η X_i αντιστοιχεί στο i-οστό αποτέλεσμα και παίρ-

νει την τιμή 1 εάν το αποτέλεσμα αυτό ανήκει στο γεγονός A. Η τιμή της τυχαίας μεταβλητής

$$X = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

είναι η προσέγγιση της $\mathrm{P}(A)$ που δίνεται από την προσομοίωση. Σύμφωνα με το Παράδειγμα 2.21, η X έχει μέση τιμή $\mathrm{P}(A)$ και διασπορά $\mathrm{P}(A)\big(1 - \mathrm{P}(A)\big)/n$, έτσι ώστε για μεγάλα n, δίνει μία ακριβή προσέγγιση της $\mathrm{P}(A)$.

2.8 ΣΥΝΟΨΗ ΚΑΙ ΣΥΖΗΤΗΣΗ

Οι τυχαίες μεταβλητές παρέχουν τα φυσικά εργαλεία για το χειρισμό μοντέλων πιθανοτήτων στα οποία το αποτέλεσμα συνδέεται με κάποιες αριθμητικές τιμές που μας ενδιαφέρουν. Στο κεφάλαιο αυτό εστιάσαμε σε διακριτές τυχαίες μεταβλητές και αναπτύξαμε ένα πλαίσιο εννοιών και σχετικά εργαλεία.

Συγκεκριμένα, εισαγάγαμε έννοιες όπως η ΣΜΠ, η μέση τιμή και η διασπορά, οι οποίες περιγράφουν με διάφορους βαθμούς λεπτομέρειας τον πιθανοτικό χαρακτήρα μίας διακριτής τυχαίας μεταβλητής. Δείξαμε τον τρόπο χρήσης της ΣΜΠ μίας τυχαίας μεταβλητής X για τον υπολογισμό της μέσης τιμής και διασποράς μίας συναφούς τυχαίας μεταβλητής $Y = g(X)$ χωρίς να υπολογίσουμε τη ΣΜΠ της Y. Στην ειδική περίπτωση όπου η g είναι μία γραμμική συνάρτηση, $Y = aX + b$, οι μέσες τιμές και οι διασπορές των X και Y σχετίζονται ως εξής:

$$\mathbf{E}[Y] = a\mathbf{E}[X] + b, \qquad \mathrm{var}(Y) = a^2\mathrm{var}(X).$$

Επίσης, συζητήσαμε μερικές ιδιαίτερες τυχαίες μεταβλητές και υπολογίσαμε τις ΣΜΠ, τις μέσες τιμές και τις διασπορές τους, όπως συνοψίζονται στον πίνακα που ακολουθεί.

Σύνοψη Αποτελεσμάτων Ειδικών Τυχαίων Μεταβλητών

Διακριτή ομοιόμορφη στο $[a, b]$:

$$p_X(k) = \begin{cases} \dfrac{1}{b-a+1}, & \text{εάν } k = a, a+1, \ldots, b, \\ 0, & \text{αλλιώς,} \end{cases}$$

$$\mathbf{E}[X] = \frac{a+b}{2}, \qquad \mathrm{var}(X) = \frac{(b-a)(b-a+2)}{12}.$$

Bernoulli με Παράμετρο p: (Περιγράφει την επιτυχία ή αποτυχία σε μία

επανάληψη.)

$$p_X(k) = \begin{cases} p, & \text{εάν } k = 1, \\ 1-p, & \text{έαν } k = 0, \end{cases}$$

$$\mathbf{E}[X] = p, \qquad \text{var}(X) = p(1-p).$$

Διωνυμική με Παραμέτρους p και n: (Περιγράφει τον αριθμό των επιτυχιών σε n ανεξάρτητες επαναλήψεις Bernoulli.)

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

$$\mathbf{E}[X] = np, \qquad \text{var}(X) = np(1-p).$$

Γεωμετρική με Παράμετρο p: (Περιγράφει τον αριθμό επαναλήψεων μέχρι την πρώτη επιτυχία, σε μία ακολουθία ανεξάρτητων επαναλήψεων Bernoulli.)

$$p_X(k) = (1-p)^{k-1} p, \qquad k = 1, 2, \ldots,$$

$$\mathbf{E}[X] = \frac{1}{p}, \qquad \text{var}(X) = \frac{1-p}{p^2}.$$

Poisson με παράμετρο λ: (Προσεγγίζει τη διωνυμική ΣΜΠ όταν το n είναι μεγάλο, το p είναι μικρό και $\lambda = np$.)

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \qquad k = 0, 1, \ldots,$$

$$\mathbf{E}[X] = \lambda, \qquad \text{var}(X) = \lambda.$$

Επίσης, θεωρήσαμε πολλαπλές τυχαίες μεταβλητές και εισαγάγαμε τις από κοινού και δεσμευμένες ΣΜΠ, και τις σχετικές μέσες τιμές τους. Οι δεσμευμένες ΣΜΠ είναι συχνά η αφετηρία μοντέλων πιθανοτήτων. Μπορούν να χρησιμοποιηθούν για τον υπολογισμό άλλων ποσοτήτων που μας ενδιαφέρουν, όπως οι οριακές ή οι από κοινού ΣΜΠ και οι μέσες τιμές, μέσω μίας ακολουθιακής μεθόδου ή μίας τακτικής "διαίρει και βασίλευε". Συγκεκριμένα, δεδομένης της υπό δέσμευση ΣΜΠ $p_{X|Y}(x \,|\, y)$:

(α) Η από κοινού ΣΜΠ μπορεί να υπολογιστεί ως

$$p_{X,Y}(x, y) = p_Y(y) p_{X|Y}(x \,|\, y).$$

Αυτό επεκτείνεται στην περίπτωση τριών η περισσοτέρων τυχαίων μεταβλη-

τών, όπως

$$p_{X,Y,Z}(x,y,z) = p_Z(z)p_{Y|Z}(y\,|\,z)p_{X|Y,Z}(x\,|\,y,z),$$

και είναι ανάλογη με την ακολουθιακή μέθοδο βασισμένη σε αναπαράσταση δένδρου χρησιμοποιώντας τον κανόνα του πολλαπλασιασμού, όπως έχει συζητηθεί στο Κεφάλαιο 1.

(β) Η οριακή ΣΜΠ μπορεί να υπολογιστεί ως

$$p_X(x) = \sum_y p_Y(y)p_{X|Y}(x\,|\,y),$$

η οποία γενικεύει την υπολογιστική μέθοδο "διαίρει και βασίλευε" όπως συζητήθηκε στο Κεφάλαιο 1.

(γ) Η υπολογιστική μέθοδος "διαίρει και βασίλευε" στο (β) παραπάνω μπορεί να επεκταθεί για τον υπολογισμό μέσων τιμών χρησιμοποιώντας το θεώρημα της συνολικής μέσης τιμής:

$$\mathbf{E}[X] = \sum_y p_Y(y)\mathbf{E}[X\,|\,Y=y].$$

Εισαγάγαμε την έννοια των ανεξάρτητων τυχαίων μεταβλητών, σε αναλογία με την έννοια των ανεξάρτητων γεγονότων. Μεταξύ άλλων θεμάτων, εστιάσαμε σε τυχαίες μεταβλητές X οι οποίες προκύπτουν προσθέτοντας ανεξάρτητες τυχαίες μεταβλητές $X_1, \ldots, X_n$:

$$X = X_1 + \cdots + X_n.$$

Αποδείξαμε ότι η μέση τιμή και η διασπορά του αθροίσματος είναι ίσες με το άθροισμα των μέσων τιμών και των διασπορών αντίστοιχα:

$$\mathbf{E}[X] = \mathbf{E}[X_1] + \cdots + \mathbf{E}[X_n], \qquad \text{var}(X) = \text{var}(X_1) + \cdots + \text{var}(X_n).$$

Ο τύπος της μέσης τιμής δεν προϋποθέτει ανεξαρτησία των X_i, ενώ ο τύπος της διασποράς το απαιτεί.

Οι έννοιες και οι μέθοδοι του κεφαλαίου αυτού επεκτείνονται κατάλληλα σε γενικές τυχαίες μεταβλητές (βλ. στο επόμενο κεφάλαιο), και είναι θεμελιώδεις για το αντικείμενο μας.

Π Ρ Ο Β Λ Η Μ Α Τ Α

ΠΑΡΑΓΡΑΦΟΣ 2.2. Συναρτήσεις Μάζας Πιθανότητας

Πρόβλημα 1. Η ομάδα ποδοσφαίρου του MIT έχει προγραμματίσει 2 παιχνίδια για το Σαββατοκύριακο. Έχει πιθανότητα 0.4 να μη χάσει στο πρώτο παιχνίδι και πιθανότητα 0.7 να μη χάσει στο δεύτερο παιχνίδι, ανεξάρτητα από το πρώτο. Εάν δεν χάσει ένα συγκεκριμένο παιχνίδι, η ομάδα έχει την ίδια πιθανότητα να κερδίσει ή να φέρει ισοπαλία, ανεξάρτητα από ό,τι συμβαίνει στο άλλο παιχνίδι. Η ομάδα θα πάρει 2 πόντους για μια νίκη, 1 για ισοπαλία και 0 για ήττα. Βρείτε τη ΣΜΠ του αριθμού των πόντων που η ομάδα θα πάρει στο Σαββατοκύριακο.

Πρόβλημα 2. Βρίσκεστε σε ένα πάρτυ με 500 καλεσμένους. Ποια είναι η πιθανότητα ακριβώς ένας άλλος καλεσμένος να έχει την ίδια μέρα γενέθλια όπως εσείς; Υπολογίστε την με ακρίβεια και επίσης προσεγγιστικά χρησιμοποιώντας την Poisson ΣΜΠ. (Για χάρη απλότητας, εξαιρέστε γενέθλια στις 29 Φεβρουαρίου.)

Πρόβλημα 3. Ο Γιάννης και ο Βασίλης παίζουν έναν αγώνα σκακιού στο οποίο ο πρώτος παίκτης που κερδίζει ένα παιχνίδι κερδίζει τον αγώνα. Μετά από 10 διαδοχικές ισοπαλίες, το αποτέλεσμα του αγώνα είναι ισοπαλία. Ο Γιάννης κερδίζει ένα παιχνίδι με πιθανότητα 0.4, ο Βασίλης με πιθανότητα 0.3, και έρχονται σε ισοπαλία με πιθανότητα 0.3, ανεξάρτητα από προηγούμενα παιχνίδια.

(α) Ποια είναι η πιθανότητα ότι ο Γιάννης κερδίζει τον αγώνα;

(β) Ποια είναι η ΣΜΠ της διάρκειας του αγώνα;

Πρόβλημα 4. Ένας παροχέας Internet χρησιμοποιεί 50 modem για να εξυπηρετήσει τις ανάγκες 1000 πελατών. Υπολογίζεται ότι σε δεδομένο χρόνο, κάθε πελάτης χρειάζεται σύνδεση με πιθανότητα 0.01, ανεξάρτητα από άλλους πελάτες.

(α) Ποια είναι η ΣΜΠ του αριθμού των modem που χρησιμοποιούνται σε μία δεδομένη χρονική στιγμή;

(β) Επαναλάβετε το μέρος (α) προσεγγίζοντας τη ΣΜΠ του αριθμού των πελατών που χρειάζονται σύνδεση με Poisson ΣΜΠ.

(γ) Ποια είναι η πιθανότητα ο αριθμός των πελατών που χρειάζονται σύνδεση να είναι μεγαλύτερος από τον αριθμό των modem; Δώστε έναν ακριβή και έναν προσεγγιστικό τύπο που βασίζεται στην προσέγγιση Poisson του μέρους (β).

Πρόβλημα 5. Ένα σύστημα επικοινωνίας πακέτων αποτελείται από ένα ενταμιευτή ο οποίος αποθηκεύει τα πακέτα που έρχονται από κάποια πηγή και στη συνέχεια τα στέλνει σε ένα δέκτη. Το σύστημα λειτουργεί σε ζεύγη χρονικών σχισμών. Στην πρώτη χρονική σχισμή, το σύστημα αποθηκεύει έναν αριθμό από πακέτα τα οποία έρχονται από την πηγή σύμφωνα με μία Poisson ΣΜΠ με παράμετρο λ· ωστόσο, ο μέγιστος αριθμός πακέτων που μπορεί να αποθηκευτεί είναι ένας δεδομένος ακέραιος b και τα πακέτα που φτάνουν σε ένα γεμάτο ενταμιευτή απορρίπτονται. Στη δεύτερη χρονική σχισμή, το σύστημα στέλνει στο δέκτη όλα τα αποθηκευμένα πακέτα ή c πακέτα (οτιδήποτε είναι μικρότερο). Εδώ ο c είναι ένας δεδομένος ακέραιος με $0 < c < b$.

(α) Αν υποθέσουμε ότι στην αρχή της πρώτης σχισμής ο ενταμιευτής είναι άδειος, βρείτε την ΣΜΠ του αριθμού των πακέτων που είναι αποθηκευμένα στο τέλος της πρώτης και στο τέλος της δεύτερης σχισμής.

(β) Ποια είναι η πιθανότητα ότι κάποια πακέτα απορρίπτονται κατά τη διάρκεια της πρώτης σχισμής;

Πρόβλημα 6. Ο Άρης και ο ΠΑΟ θα παίξουν μία σειρά n παιχνιδιών μπάσκετ, όπου το n είναι περιττός. Ο Άρης έχει πιθανότητα p να κερδίσει οποιοδήποτε παιχνίδι, ανεξάρτητα από άλλα παιχνίδια.

(α) Βρείτε τις τιμές της p για τις οποίες μία σειρά με $n = 5$ παιχνίδια είναι προτιμώτερη για τον Άρη από μία σειρά με $n = 3$.

(β) Γενικεύστε το μέρος (α), δηλαδή για οποιοδήποτε $k > 0$, βρείτε τις τιμές της p για τις οποίες το $n = 2k + 1$ είναι προτιμώτερο για τον Άρη από το $n = 2k - 1$.

Πρόβλημα 7. Μόλις νοικιάσατε ένα μεγάλο σπίτι και ο μεσίτης σας έδωσε 5 κλειδιά, ένα για κάθε μία από τις 5 πόρτες του σπιτιού. Δυστυχώς, όλα τα κλειδιά φαίνονται ίδια και για να ανοίξετε την εξώπορτα τα δοκιμάζετε τυχαία.

(α) Βρείτε τη ΣΜΠ του αριθμού των προσπαθειών που θα χρειαστείτε για να ανοίξετε την πόρτα, σύμφωνα με τις παρακάτω εναλλακτικές υποθέσεις: (1) μετά από μία αποτυχημένη προσπάθεια, σημειώνετε το αντίστοιχο κλειδί, έτσι ώστε να μην το ξαναχρησιμοποιήσετε και (2) σε κάθε προσπάθεια έχετε την ίδια πιθανότητα να επιλέξετε οποιοδήποτε κλειδί.

(β) Επαναλάβετε το μέρος (α) για την περίπτωση όπου ο μεσίτης σας έδωσε από δύο κλειδιά για καθεμία από τις 5 πόρτες.

Πρόβλημα 8. Αναδρομικός υπολογισμός της διωνυμικής ΣΜΠ. Έστω ότι X είναι μία διωνυμική τυχαία μεταβλητή με παραμέτρους n και p. Δείξτε ότι η ΣΜΠ της μπορεί να υπολογιστεί αρχίζοντας με $p_X(0) = (1-p)^n$ και χρησιμοποιώντας τον επαγωγικό τύπο

$$p_X(k+1) = \frac{p}{1-p} \cdot \frac{n-k}{k+1} \cdot p_X(k), \qquad k = 0, 1, \ldots, n-1.$$

Πρόβλημα 9. Το σχήμα της διωνυμικής ΣΜΠ. Θεωρήστε μία διωνυμική τυχαία μεταβλητή X με παραμέτρους n και p. Έστω k^* ο μέγιστος ακέραιος ο οποίος είναι μικρότερος ή ίσος του $(n+1)p$. Δείξτε ότι η ΣΜΠ $p_X(k)$ είναι μονοτονικά μη φθίνουσα ως προς k στο διάστημα από 0 έως k^* και είναι μονοτονικά φθίνουσα ως προς k για $k \geq k^*$.

Πρόβλημα 10. Το σχήμα της Poisson ΣΜΠ. Έστω ότι X είναι μία τυχαία μεταβλητή Poisson με παράμετρο λ. Δείξτε ότι η ΣΜΠ $p_X(k)$ αυξάνει μονοτονικά ως προς k μέχρι το σημείο όπου το k ισούται με το μέγιστο ακέραιο που δεν ξεπερνά την λ και μετά το σημείο αυτό είναι μονοτονικά φθίνουσα ως προς k.

Πρόβλημα 11.* Το πρόβλημα του σπιρτόκουτου – εμπνευσμένο από τις συνήθειες του Banach. Ένας μαθηματικός που είναι καπνιστής έχει ένα σπιρτόκουτο στη δεξιά του τσέπη και ένα στην αριστερή του τσέπη. Κάθε φορά που θέλει να ανάψει ένα τσιγάρο, επιλέγει ένα σπιρτόκουτο από οποιαδήποτε τσέπη του με πιθανότητα $p = 1/2$, ανεξάρτητα από προηγούμενες επιλογές. Τα δύο σπιρτόκουτα έχουν αρχικά n σπίρτα το καθένα. Ποια είναι η ΣΜΠ του αριθμού των σπίρτων που απομένουν τη στιγμή που ο μαθηματικός ψάχνει για σπίρτο και ανακαλύπτει ότι το αντίστοιχο σπιρτόκουτο είναι άδειο; Πως μπορούμε να γενικεύσουμε στην περίπτωση όπου οι πιθανότητες επιλογής της αριστερής ή της δεξιάς τσέπης είναι p και $1-p$, αντίστοιχα;

Λύση. Έστω X ο αριθμός των σπίρτων που απομένουν όταν ένα από τα σπιρτόκουτα είναι άδειο. Για $k = 0, 1, \ldots, n$, έστω ότι L_k (ή R_k) είναι το γεγονός ότι ένα άδειο κουτί ανακαλύπτεται πρώτα στην αριστερή (αντίστοιχα, δεξιά) τσέπη, ενώ ο αριθμός των σπίρτων στη δεξιά (αντίστοιχα, αριστερή) τσέπη είναι k κάθε φορά. Η ΣΜΠ της X είναι

$$p_X(k) = \mathrm{P}(L_k) + \mathrm{P}(R_k), \qquad k = 0, 1, \ldots, n.$$

Αν θεωρήσουμε την επιλογή της αριστερής και της δεξιάς τσέπης σαν μία "επιτυχία" και μία "αποτυχία", αντίστοιχα, η $\mathrm{P}(L_k)$ είναι η πιθανότητα ότι υπάρχουν n επιτυχίες στις πρώτες $2n-k$ επαναλήψεις και η επανάληψη $2n-k+1$ είναι μία επιτυχία, άρα

$$\mathrm{P}(L_k) = \frac{1}{2}\binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}, \qquad k = 0, 1, \ldots, n.$$

Λόγω συμμετρίας, $\mathrm{P}(L_k) = \mathrm{P}(R_k)$, άρα

$$p_X(k) = \mathrm{P}(L_k) + \mathrm{P}(R_k) = \binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}, \qquad k = 0, 1, \ldots, n.$$

Στην πιο γενική περίπτωση, όπου οι πιθανότητες της αριστερής και δεξιάς επιλογής είναι p και $1-p$, παρόμοια έχουμε

$$\mathrm{P}(L_k) = p\binom{2n-k}{n}p^n(1-p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

και

$$\mathrm{P}(R_k) = (1-p)\binom{2n-k}{n} p^{n-k}(1-p)^n, \qquad k = 0, 1, \ldots, n,$$

το οποίο δίνει

$$\begin{aligned} p_X(k) &= \mathrm{P}(L_k) + \mathrm{P}(R_k) \\ &= \binom{2n-k}{n}\big(p^{n+1}(1-p)^{n-k} + p^{n-k}(1-p)^{n+1}\big), \qquad k = 0, 1, \ldots, n. \end{aligned}$$

Πρόβλημα 12.* Απόδειξη της προσεγγιστικής ιδιότητας της Poisson. Θεωρήστε τη ΣΜΠ μίας διωνυμικής τυχαίας μεταβλητής με παραμέτρους n και p. Δείξτε ότι ασυμπτωτικά καθώς

$$n \to \infty, \qquad p \to 0,$$

και συγχρόνως η np είναι σταθερή σε κάποια δεδομένη τιμή λ, η ΣΜΠ αυτή πλησιάζει τη ΣΜΠ της τυχαίας μεταβλητής Poisson με παράμετρο λ.

Λύση. Χρησιμοποιώντας την εξίσωση $\lambda = np$, γράφουμε τη διωνυμική ΣΜΠ ως

$$\begin{aligned} p_X(k) &= \frac{n!}{(n-k)!\,k!} p^k (1-p)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}. \end{aligned}$$

Κρατάμε το k σταθερό και αφήνουμε το $n \to \infty$. Έχουμε, για $j = 1, \ldots, k$,

$$\frac{n-k+j}{n} \to 1, \quad \left(1 - \frac{\lambda}{n}\right)^{-k} \to 1, \quad \left(1 - \frac{\lambda}{n}\right)^{n} \to e^{-\lambda}.$$

Άρα για κάθε συγκεκριμένο k, καθώς το $n \to \infty$ έχουμε

$$p_X(k) \to e^{-\lambda}\frac{\lambda^k}{k!}.$$

ΠΑΡΑΓΡΑΦΟΣ 2.3. Συναρτήσεις Τυχαίων Μεταβλητών

Πρόβλημα 13. Μία οικογένεια έχει 5 φυσικά παιδιά και έχει υιοθετήσει 2 κορίτσια. Κάθε φυσικό παιδί έχει ίση πιθανότητα να είναι αγόρι ή κορίτσι, ανεξάρτητα από τα άλλα παιδιά. Για τα 7 παιδιά βρείτε τη ΣΜΠ του αριθμού των κοριτσιών.

Πρόβλημα 14. Έστω X μία τυχαία μεταβλητή η οποία παίρνει τιμές από 0 μέχρι 9 με ίση πιθανότητα $1/10$.

(α) Βρείτε τη ΣΜΠ της τυχαίας μεταβλητής $Y = X \bmod(3)$.

(β) Βρείτε τη ΣΜΠ της τυχαίας μεταβλητής $Y = 5 \bmod(\mathrm{X}+1)$.

Πρόβλημα 15. Έστω K μία τυχαία μεταβλητή η οποία παίρνει ακέραιες τιμές στο διάστημα $[-n, n]$, με ίση πιθανότητα $1/(2n+1)$. Βρείτε τη ΣΜΠ της τυχαίας μεταβλητής $Y = \ln X$, όπου $X = a^{|K|}$ και $a > 0$.

ΠΑΡΑΓΡΑΦΟΣ 2.4. Μέση Τιμή και Διασπορά

Πρόβλημα 16. Έστω X μία τυχαία μεταβλητή με ΣΜΠ

$$p_X(x) = \begin{cases} x^2/a, & \text{εάν } x = -3, -2, -1, 0, 1, 2, 3, \\ 0, & \text{διαφορετικά.} \end{cases}$$

(α) Βρείτε την τιμή του a και την $\mathbf{E}[X]$.

(β) Ποια είναι η ΣΜΠ της τυχαίας μεταβλητής $Z = \big(X - \mathbf{E}[X]\big)^2$;

(γ) Υπολογίστε τη διασπορά της X, χρησιμοποιώντας το μέρος (β).

(δ) Υπολογίστε τη διασπορά της X χρησιμοποιώντας τον τύπο $\mathrm{var}(X) = \sum_x \big(x - \mathbf{E}[X]\big)^2 p_X(x)$.

Πρόβλημα 17. Η θερμοκρασία μίας πόλης μοντελοποιείται ως μία τυχαία μεταβλητή με μέση τιμή και τυπική απόκλιση ίσες με 10 βαθμούς Κελσίου. Μία ημέρα θεωρείται "κανονική" εάν η θερμοκρασία κατά τη διάρκειά της κυμαίνεται εντός μίας τυπικής απόκλισης από τη μέση τιμή. Ποια θα ήταν η διακύμανση της θερμοκρασίας για μία κανονική ημέρα εάν η θερμοκρασία εκφραζόταν σε βαθμούς Φαρενάιτ;

Πρόβλημα 18. Έστω a και b θετικοί ακέραιοι με $a \leq b$ και έστω X μία τυχαία μεταβλητή η οποία παίρνει ως τιμές τις δυνάμεις του 2 στο διάστημα $[2^a, 2^b]$, με ίση πιθανότητα. Βρείτε τη μέση τιμή και τη διασπορά της X.

Πρόβλημα 19. Ένα δώρο τοποθετείται τυχαία σε ένα από 10 κουτιά, τα οποία είναι αριθμημένα από το 1 μέχρι το 10. Ψάχνετε για το δώρο κάνοντας ερωτήσεις που απαντώνται με ναι ή όχι. Βρείτε τη μέση τιμή του αριθμού των ερωτήσεων μέχρι να σιγουρευτείτε για το κουτί στο οποίο βρίσκεται το δώρο, σύμφωνα με τις παρακάτω στρατηγικές:

(α) Στρατηγική αρίθμησης: κάνοντας ερωτήσεις της μορφής "είναι στο κουτί k;".

(β) Στρατηγική διχοτόμησης: αποκλείετε περίπου μισά από τα υπολειπόμενα κουτιά κάνοντας ερωτήσεις της μορφής "είναι σε κουτί που έχει αριθμό μικρότερο ή ίσο του k;".

Πρόβλημα 20. Σε μία διαφημιστική εκστρατεία, μία βιομηχανία σοκολάτας τοποθετεί χρυσά κουπόνια σε μερικές από τις σοκολάτες της, με την υπόσχεση ότι ένα χρυσό κουπόνι σας παρέχει όλη τη σοκολάτα που μπορείτε να φάτε σε όλης σας τη ζωή. Εάν η πιθανότητα εύρεσης ενός κουπονιού είναι p, βρείτε την μέση τιμή και διασπορά του αριθμού από σοκολάτες που χρειάζεται να αγοράσετε για να βρείτε ένα κουπόνι.

Πρόβλημα 21. Το παράδοξο της Αγίας Πετρούπολης. Ρίχνετε ανεξάρτητα ένα αμερόληπτο νόμισμα και μετράτε τον αριθμό των ρίψεων μέχρι να εμφανιστούν γράμματα. Εάν ο αριθμός αυτός είναι n, λαμβάνετε 2^n ευρώ. Ποια είναι η μέση τιμή του ποσού που θα λάβετε; Πόσο θα πληρώνατε για να παίξετε το παιχνίδι αυτό;

Πρόβλημα 22. Ρίχνουμε δύο νομίσματα συγχρόνως μέχρι το ένα από αυτά να έρθει κορώνα και το άλλο γράμματα. Το πρώτο νόμισμα έρχεται κορώνα με πιθανότητα p και το δεύτερο με πιθανότητα q. Όλες οι ρίψεις θεωρούνται ανεξάρτητες.

(α) Βρείτε τη ΣΜΠ, τη μέση τιμή και τη διασπορά του αριθμού των ρίψεων.

(β) Ποια είναι η πιθανότητα ότι η τελευταία ρίψη του πρώτου νομίσματος προκύπτει κορώνα;

Πρόβλημα 23.

(α) Ρίχνουμε ένα αμερόληπτο νόμισμα συνεχώς μέχρι να παρουσιαστούν δύο διαδοχικές κορώνες. Βρείτε τη ΣΜΠ, τη μέση τιμή και τη διασπορά του αριθμού των ρίψεων.

(β) Υποθέστε τώρα ότι ρίχνουμε το νόμισμα συνεχώς μέχρι να έρθουν γράμματα αμέσως μετά από κορώνα. Βρείτε τη ΣΜΠ και τη μέση τιμή του αριθμού των ρίψεων.

Πρόβλημα 24.* Διασπορά Poisson. Θεωρήστε την τυχαία μεταβλητή Poisson με παράμετρο λ. Υπολογίστε τη δεύτερη ροπή και τη διασπορά της.

Λύση. Όπως έχουμε δείξει στο κείμενο, η μέση τιμή είναι

$$\mathbf{E}[X] = \lambda.$$

Επίσης

$$\begin{aligned}
\mathbf{E}[X^2] &= \sum_{k=1}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \lambda \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}
\end{aligned}$$

$$
\begin{aligned}
&= \lambda \sum_{m=0}^{\infty} (m+1) \frac{e^{-\lambda} \lambda^m}{m!} \\
&= \lambda \big(\mathbf{E}[X] + 1 \big) \\
&= \lambda(\lambda + 1),
\end{aligned}
$$

από το οποίο έχουμε

$$\mathrm{var}(X) = \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2 = \lambda(\lambda+1) - \lambda^2 = \lambda.$$

ΠΑΡΑΓΡΑΦΟΣ 2.5. Από Κοινού ΣΜΠ Πολλαπλών Τυχαίων Μεταβλητών

Πρόβλημα 25. Ένας χρηματιστής αγοράζει 100 μετοχές Α και 200 μετοχές Β. Έστω X και Y οι αλλαγές στις τιμές των Α και Β, αντίστοιχα, σε μία δεδομένη χρονική περίοδο και υποθέστε ότι η από κοινού ΣΜΠ των X και Y είναι ομοιόμορφη στο σύνολο των ακεραίων x και y οι οποίοι ικανοποιούν τις σχέσεις

$$-2 \leq x \leq 4, \qquad -1 \leq y - x \leq 1.$$

(α) Βρείτε τις οριακές ΣΜΠ και τις μέσες τιμές των X και Y.

(β) Βρείτε τη μέση τιμή του κέρδους του χρηματιστή.

Πρόβλημα 26. Μία τάξη από n φοιτητές εξετάζεται σε ένα διαγώνισμα που έχει m ερωτήσεις. Υποθέστε ότι ο φοιτητής i δίνει απαντήσεις στις πρώτες m_i ερωτήσεις.

(α) Ο βαθμολογητής τυχαία επιλέγει μία απάντηση, ας την πούμε (I, J), όπου I είναι ο αριθμός μητρώου του φοιτητή (που παίρνει τιμές $1, \ldots, n$) και J είναι ο αριθμός της ερώτησης (που παίρνει τιμές $1, \ldots, m$). Υποθέστε ότι όλες οι απαντήσεις επιλέγονται με ίση πιθανότητα. Υπολογίστε τις οριακές και τις από κοινού ΣΜΠ των I και J.

(β) Υποθέστε ότι η απάντηση στην ερώτηση j που δόθηκε από το φοιτητή i είναι σωστή με πιθανότητα p_{ij}. Κάθε απάντηση παίρνει a πόντους εάν είναι σωστή και b πόντους διαφορετικά. Υπολογίστε τη μέση τιμή της βαθμολογίας του φοιτητή i.

Πρόβλημα 27. Η ΣΜΠ του ελαχίστου μερικών τυχαίων μεταβλητών. Σε μία δεδομένη μέρα, το σκορ σας στο γκολφ παίρνει τιμές στο διάστημα από 101 μέχρι 110, με πιθανότητα 0.1, ανεξάρτητα από άλλες μέρες. Για να βελτιώσετε το σκορ σας, αποφασίζετε να παίξετε σε τρεις διαφορετικές μέρες και να δηλώσετε ως σκορ σας το ελάχιστο X των σκορ X_1, X_2 και X_3 στις διαφορετικές μέρες.

(α) Υπολογίστε τη ΣΜΠ της X.

(β) Πόσο έχει βελτιωθεί η μέση τιμή του σκορ σας σαν αποτέλεσμα του να παίξετε στις τρεις μέρες;

Πρόβλημα 28.* Το πρόβλημα του παιχνιδιού γνώσεων. Θεωρήστε ένα διαγώνισμα όπου ένα άτομο έχει μία λίστα από n ερωτήσεις και μπορεί να τις απαντήσει με οποιαδήποτε σειρά επιλέξει. Η ερώτηση i θα απαντηθεί σωστά με πιθανότητα p_i και το άτομο τότε θα πάρει μία αμοιβή v_i. Στην πρώτη λανθασμένη απάντηση, το διαγώνισμα τελειώνει και το άτομο θα κρατήσει τις προηγούμενες αμοιβές του. Το πρόβλημα είναι να επιλέξετε τη σειρά των ερωτήσεων, έτσι ώστε να μεγιστοποιηθεί η μέση τιμή της συνολικής αμοιβής. Δείξτε ότι είναι βέλτιστο να απαντήσει κανείς τις ερωτήσεις σύμφωνα με την μη αύξουσα σειρά των $p_i v_i/(1-p_i)$.

Λύση. Θα χρησιμοποιήσουμε το λεγόμενο επιχείρημα ανταλλαγής που χρησιμοποιείται συχνά σε προβλήματα χρονοπρογραμματισμού. Έστω i και j η k-οστή ερώτηση και η $(k+1)$-οστή ερωτήση σε μία βέλτιστα διατεταγμένη μετάθεση

$$L = (i_1, \ldots, i_{k-1}, i, j, i_{k+2}, \ldots, i_n).$$

Θεωρήστε τη μετάθεση

$$L' = (i_1, \ldots, i_{k-1}, j, i, i_{k+2}, \ldots, i_n)$$

που λαμβάνεται από την L ανταλλάσσοντας τη διάταξη των ερωτήσεων i και j. Υπολογίζουμε τις μέσες τιμές των αμοιβών της L και L', και χρησιμοποιούμε το γεγονός ότι εφόσον η L είναι βέλτιστα διατεταγμένη, έχουμε

$$\mathbf{E}[\text{αμοιβή στην } L] \geq \mathbf{E}[\text{αμοιβή στην } L'].$$

Ορίζουμε το *βάρος* της ερώτησης i να είναι

$$w(i) = \frac{p_i v_i}{(1-p_i)}.$$

Θα δείξουμε ότι οποιαδήποτε μετάθεση των ερωτήσεων σύμφωνα με μη αύξουσα διάταξη των βαρών τους, μεγιστοποιεί τη μέση αμοιβή.

Εάν $L = (i_1, \ldots, i_n)$ είναι μία μετάθεση των ερωτήσεων, ορίζουμε την $L^{(k)}$ ως τη μετάθεση που προκύπτει από την L ανταλλάσοντας τις ερωτήσεις i_k και i_{k+1}. Ας υπολογίσουμε κατ' αρχάς τη διαφορά μεταξύ της μέσης αμοιβής της L και της $L^{(k)}$. Έχουμε

$$\mathbf{E}[\text{αμοιβή της } L] = p_{i_1} v_{i_1} + p_{i_1} p_{i_2} v_{i_2} + \cdots + p_{i_1} \cdots p_{i_n} v_{i_n},$$

και

$$\begin{aligned}\mathbf{E}\big[\text{αμοιβή της } L^{(k)}\big] = & \, p_{i_1}v_{i_1} + p_{i_1}p_{i_2}v_{i_2} + \cdots + p_{i_1}\cdots p_{i_{k-1}}v_{i_{k-1}} \\ & + p_{i_1}\cdots p_{i_{k-1}}p_{i_{k+1}}v_{i_{k+1}} + p_{i_1}\cdots p_{i_{k-1}}p_{i_{k+1}}p_{i_k}v_{i_k} \\ & + p_{i_1}\cdots p_{i_{k+2}}v_{i_{k+2}} + \cdots + p_{i_1}\cdots p_{i_n}v_{i_n}.\end{aligned}$$

Άρα

$$\begin{aligned}\mathbf{E}\big[\text{αμοιβή της } L^{(k)}\big] - \mathbf{E}[\text{αμοιβή της } L] &= p_{i_1}\cdots p_{i_{k-1}}\big(p_{i_{k+1}}v_{i_{k+1}} + p_{i_{k+1}}p_{i_k}v_{i_k} \\ &\qquad - p_{i_k}v_{i_k} - p_{i_k}p_{i_{k+1}}v_{i_{k+1}}\big) \\ &= p_{i_1}\cdots p_{i_{k-1}}(1-p_{i_k})(1-p_{i_{k+1}})\big(w(i_{k+1}) - w(i_k)\big).\end{aligned}$$

Τώρα, ας επιστρέψουμε στο πρόβλημά μας. Θεωρήστε οποιαδήποτε μετάθεση L των ερωτήσεων. Εάν $w(i_k) < w(i_{k+1})$ για κάποιο k, έπεται από την παραπάνω σχέση ότι η μετάθεση $L^{(k)}$ έχει μέση αμοιβή μεγαλύτερη από αυτή της L. Άρα, μία βέλτιστη μετάθεση των ερωτήσεων πρέπει να είναι σύμφωνα με μία αύξουσα διάταξη των βαρών τους.

Τελικά θα δείξουμε ότι οποιεσδήποτε δύο τέτοιες μεταθέσεις, έχουν ίσες μέσες αμοιβές. Υποθέστε ότι η L είναι μία τέτοια μετάθεση και έστω $w(i_k) = w(i_{k+1})$ για κάποιο k. Γνωρίζουμε ότι ανταλλάσσοντας τη διάταξη των i_k και i_{k+1} διατηρείται η μέση αμοιβή. Άρα, η μέση αμοιβή οποιασδήποτε μετάθεσης L' βασισμένης σε μία μη αύξουσα διάταξη των βαρών ισούται με αυτή της L, διότι η L' λαμβάνεται από την L ανταλλάσσοντας επανειλημμένα παρακείμενες ερωτήσεις που έχουν ίσα βάρη.

Πρόβλημα 29.* Η σχέση εγκλεισμού-αποκλεισμού. Έστω $A_1, A_2, \ldots, A_n$ γεγονότα. Έστω $S_1 = \{i \,|\, 1 \le i \le n\}$, $S_2 = \{(i_1, i_2) \,|\, 1 \le i_1 < i_2 \le n\}$, και πιο γενικά, έστω ότι S_m είναι το σύνολο όλων των m-άδων $(i_1, \ldots, i_m)$ των δεικτών που ικανοποιούν $1 \le i_1 < i_2 < \cdots < i_m \le n$. Δείξτε ότι

$$\begin{aligned}\mathrm{P}\left(\cup_{k=1}^n A_k\right) = & \sum_{i \in S_1} \mathrm{P}(A_i) - \sum_{(i_1,i_2) \in S_2} \mathrm{P}(A_{i_1} \cap A_{i_2}) \\ & + \sum_{(i_1,i_2,i_3) \in S_3} \mathrm{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots + (-1)^{n-1}\mathrm{P}\left(\cap_{k=1}^n A_k\right).\end{aligned}$$

Υπόδειξη: Έστω X_i μία δυαδική τυχαία μεταβλητή η οποία ισούται με 1 όταν το A_i συμβαίνει και με 0 διαφορετικά. Συσχετίστε το γεγονός που μας ενδιαφέρει με την τυχαία μεταβλητή $(1-X_1)(1-X_2)\cdots(1-X_n)$.

Λύση. Ας εκφράσουμε το γεγονός $B = \cup_{k=1}^n A_k$ σαν συνάρτηση των τυχαίων μεταβλητών $X_1, \ldots, X_n$. Το γεγονός B^c συμβαίνει όταν όλες οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ είναι μηδέν, το οποίο συμβαίνει όταν η τυχαία μεταβλητή $Y = (1-X_1)(1-X_2)\cdots(1-X_n)$ είναι ίση με 1. Παρατηρήστε ότι η Y μπορεί να πάρει τιμές μόνο στο σύνολο $\{0, 1\}$. Άρα,

$$\mathrm{P}(B^c) = \mathrm{P}(Y = 1) = \mathbf{E}[Y].$$

Επομένως,

$$\begin{aligned}\mathrm{P}(B) &= 1 - \mathbf{E}\big[(1-X_1)(1-X_2)\cdots(1-X_n)\big] \\ &= \mathbf{E}[X_1+\cdots+X_n] - \mathbf{E}\left[\sum_{(i_1,i_2)\in S_2} X_{i_1}X_{i_2}\right] + \cdots + (-1)^{n-1}\mathbf{E}[X_1\cdots X_n].\end{aligned}$$

Παρατηρούμε ότι

$$\mathbf{E}[X_i] = \mathrm{P}(A_i), \qquad \mathbf{E}[X_{i_1}X_{i_2}] = \mathrm{P}(A_{i_1}\cap A_{i_2}),$$
$$\mathbf{E}[X_{i_1}X_{i_2}X_{i_3}] = \mathrm{P}(A_{i_1}\cap A_{i_2}\cap A_{i_3}), \qquad \mathbf{E}\big[X_1X_2\cdots X_n\big] = \mathrm{P}(\cap_{k=1}^n A_k),$$

κτλ., από τις οποίες έπεται ο επιθυμητός τύπος.

Πρόβλημα 30.* Η βάση δεδομένων των φίλων του Αριστείδη περιέχει n εγγραφές, αλλά λόγω ενός λάθους στο λογισμικό, οι διευθύνσεις αντιστοιχούν στα ονόματα με εντελώς τυχαίο τρόπο. Ο Αριστείδης γράφει μία κάρτα ευχών σε καθένα από τους φίλους του και τις στέλνει στη (πιθανώς λανθασμένη λόγω λογισμικού) διεύθυνση. Ποια είναι η πιθανότητα ότι ένας τουλάχιστον από τους φίλους του θα πάρει τη σωστή κάρτα; *Υπόδειξη:* Χρησιμοποιήστε τη σχέση εγκλεισμού-αποκλεισμού.

Λύση. Έστω ότι A_k είναι το γεγονός ότι η k-οστή κάρτα στέλνεται στη σωστή διεύθυνση. Έχουμε για οποιαδήποτε k, j, i,

$$\mathrm{P}(A_k) = \frac{1}{n} = \frac{(n-1)!}{n!},$$
$$\mathrm{P}(A_k\cap A_j) = \mathrm{P}(A_k)\mathrm{P}(A_j \mid A_k) = \frac{1}{n}\cdot\frac{1}{n-1} = \frac{(n-2)!}{n!},$$
$$\mathrm{P}(A_k\cap A_j\cap A_i) = \frac{1}{n}\cdot\frac{1}{n-1}\cdot\frac{1}{n-2} = \frac{(n-3)!}{n!},$$

κτλ. και

$$\mathrm{P}(\cap_{k=1}^n A_k) = \frac{1}{n!}.$$

Αν εφαρμόσουμε τη σχέση εγκλεισμού-αποκλεισμού,

$$\begin{aligned}\mathrm{P}\left(\cup_{k=1}^n A_k\right) &= \sum_{i\in S_1}\mathrm{P}(A_i) - \sum_{(i_1,i_2)\in S_2}\mathrm{P}(A_{i_1}\cap A_{i_2}) \\ &\quad + \sum_{(i_1,i_2,i_3)\in S_3}\mathrm{P}(A_{i_1}\cap A_{i_2}\cap A_{i_3}) - \cdots + (-1)^{n-1}\mathrm{P}\left(\cap_{k=1}^n A_k\right),\end{aligned}$$

έχουμε την επιθυμητή πιθανότητα

$$\begin{aligned}\mathrm{P}(\cup_{k=1}^{n} A_k) &= \binom{n}{1}\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \binom{n}{3}\frac{(n-3)!}{n!} - \cdots + (-1)^{n-1}\frac{1}{n!} \\ &= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!}.\end{aligned}$$

Όταν το n είναι μεγάλο, η πιθανότητα αυτή μπορεί να προσεγγιστεί από την $1 - e^{-1}$.

ΠΑΡΑΓΡΑΦΟΣ 2.6. ΔΕΣΜΕΥΣΗ

Πρόβλημα 31. Θεωρήστε τέσσερις ανεξάρτητες ρίψεις ενός 6-έδρου ζαριού. Έστω X ο αριθμός των 1 και έστω Y ο αριθμός των 2 που προκύπτουν. Ποια είναι η από κοινού ΣΜΠ των X και Y;

Πρόβλημα 32. Το πρόβλημα της από κοινού ζωής του D. Bernoulli. Θεωρήστε $2m$ άτομα που δημιουργούν m ζευγάρια τα οποία αρχικά ζουν μαζί. Υποθέστε ότι μετά από κάποιο χρονικό διάστημα, η πιθανότητα κάθε ατόμου να ζει είναι p, ανεξάρτητα από τα άλλα άτομα. Εκείνη τη στιγμή, έστω A ο αριθμός των ατόμων στα οποία ζουν και έστω S ο αριθμός των ζευγαριών στα οποία ζουν και οι δύο σύντροφοι. Για οποιοδήποτε αριθμό επιζώντων a, βρείτε την $\mathbf{E}[S \mid A = a]$.

Πρόβλημα 33.* Ένα νόμισμα έχει πιθανότητα να έρθει κορώνα ίση με p. Το ρίχνουμε διαδοχικά και ανεξάρτητα μέχρι να έρθει κορώνα δυο φορές στη σειρά, ή γράμματα δύο φορές στη σειρά. Βρείτε τη μέση τιμή του αριθμού των ρίψεων.

Λύση. Μία δυνατότητα είναι να υπολογίσουμε τη ΣΜΠ του αριθμού των ρίψεων X, μέχρι να τελειώσει το παιχνίδι και στη συνέχεια να τη χρησιμοποιήσουμε για να υπολογίσουμε την $\mathbf{E}[X]$. Ωστόσο με ένα κίβδηλο νόμισμα αυτό είναι άβολο, ώστε προτιμούμε τη χρήση του θεωρήματος της συνολικής μέσης τιμής και μίας αρμόζουσας διαμέρισης του δειγματικού χώρου. Έστω H_k (ή T_k) το γεγονός ότι έρχεται κορώνα (ή γράμματα, αντίστοιχα) στην k-οστή ρίψη και έστω p (αντίστοιχα, q) η πιθανότητα του H_k (αντίστοιχα, T_k). Εφόσον H_1 και T_1 δημιουργούν μία διαμέριση του δειγματικού χώρου και $\mathrm{P}(H_1) = p$ και $\mathrm{P}(T_1) = q$, έχουμε

$$\mathbf{E}[X] = p\mathbf{E}[X \mid H_1] + q\mathbf{E}[X \mid T_1].$$

Αν χρησιμοποιήσουμε ξανά το θεώρημα της συνολικής μέσης τιμής, έχουμε

$$\mathbf{E}[X \mid H_1] = p\mathbf{E}[X \mid H_1 \cap H_2] + q\mathbf{E}[X \mid H_1 \cap T_2] = 2p + q\big(1 + \mathbf{E}[X \mid T_1]\big),$$

όπου χρησιμοποιήσαμε τις σχέσεις

$$\mathbf{E}[X \mid H_1 \cap H_2] = 2$$

(το παιχνίδι τελειώνει μετά από δύο διαδοχικές κορώνες) και

$$\mathbf{E}[X \mid H_1 \cap T_2] = 1 + \mathbf{E}[X \mid T_1]$$

(εάν το παιχνίδι δεν τελειώνει, μόνο η τελευταία ρίψη έχει σημασία για τον προσδιορισμό του αριθμού των επιπλέον ρίψεων μέχρι τον τερματισμό). Παρόμοια, έχουμε

$$\mathbf{E}[X \mid T_1] = 2q + p\big(1 + \mathbf{E}[X \mid H_1]\big).$$

Αν συνδυάσουμε τις δύο παραπάνω σχέσεις και χρησιμοποιήσουμε το γεγονός $p+q=1$, έχουμε μετά από κάποιους υπολογισμούς

$$\mathbf{E}[X \mid T_1] = \frac{2+p^2}{1-pq},$$

και παρόμοια

$$\mathbf{E}[X \mid H_1] = \frac{2+q^2}{1-pq}.$$

Άρα,

$$\mathbf{E}[X] = p \cdot \frac{2+q^2}{1-pq} + q \cdot \frac{2+p^2}{1-pq},$$

και τελικά, χρησιμοποιώντας το γεγονός ότι $p+q=1$,

$$\mathbf{E}[X] = \frac{2+pq}{1-pq}.$$

Για την περίπτωση ενός αμερόληπτου ζαριού ($p = q = 1/2$), έχουμε $\mathbf{E}[X] = 3$. Επίσης, μπορεί να επαληθευτεί ότι $2 \leq \mathbf{E}[X] \leq 3$ για όλες τις τιμές της p.

Πρόβλημα 34.* Μία αράχνη και μία μύγα κινούνται κατά μήκος μίας ευθείας γραμμής. Σε κάθε δευτερόλεπτο, η μύγα μετακινείται κατά ένα μοναδιαίο διάστημα προς τα δεξιά ή προς τα αριστερά με την ίδια πιθανότητα p και μένει εκεί που είναι με πιθανότητα $1-2p$. Η αράχνη προχωρεί ένα μοναδιαίο διάστημα προς την κατεύθυνση της μύγας. Η αράχνη και η μύγα αρχίζουν D διαστήματα μακριά η μία από την άλλη, όπου D είναι μία τυχαία μεταβλητή η οποία παίρνει θετικές ακέραιες τιμές με δεδομένη ΣΜΠ. Εάν η αράχνη προσγειωθεί πάνω στη μύγα, η διαδικασία τελειώνει. Ποια είναι η μέση τιμή του χρόνου που χρειάζεται για να συμβεί αυτό;

Λύση. Έστω T ο χρόνος κατά τον οποίο η αράχνη προσγειώνεται πάνω στη μύγα. Ορίζουμε

A_d: το γεγονός ότι αρχικά η αράχνη και η μύγα είναι d μοναδιαία διαστήματα μακριά,

B_d: το γεγονός ότι μετά από ένα δευτερόλεπτο η αράχνη και η μύγα είναι d μοναδιαία διαστήματα μακριά.

Η προσέγγιση μας θα είναι πρώτα να εφαρμόσουμε τη δεσμευμένη μορφή του θεωρήμα-

τος της συνολικής μέσης τιμής για να υπολογίσουμε την $\mathbf{E}[T \mid A_1]$, μετά να χρησιμοποιήσουμε το αποτέλεσμα για να υπολογίσουμε την $\mathbf{E}[T \mid A_2]$ και παρόμοια να υπολογίσουμε στη σειρά τις $\mathbf{E}[T \mid A_d]$ για όλες τις τιμές του d. Τελικά, θα εφαρμόσουμε το (χωρίς δέσμευση) θεώρημα της συνολικής μέσης τιμής για να υπολογίσουμε την $\mathbf{E}[T]$.

Έχουμε

$$A_d = (A_d \cap B_d) \cup (A_d \cap B_{d-1}) \cup (A_d \cap B_{d-2}), \qquad \text{εάν } d > 1.$$

Αυτό συμβαίνει διότι εάν η αράχνη και η μύγα βρίσκονται σε απόσταση $d > 1$, τότε ένα δευτερόλεπτο αργότερα η απόσταση τους θα είναι d (εάν η μύγα απομακρυνθεί από την αράχνη) ή $d-1$ (εάν η μύγα δεν μετακινηθεί) ή $d-2$ (εάν η μύγα μετακινηθεί προς την αράχνη). Έχουμε επίσης για την περίπτωση όπου η αράχνη και η μύγα αρχίζουν ένα μοναδιαίο διάστημα μακριά η μία από την άλλη

$$A_1 = (A_1 \cap B_1) \cup (A_1 \cap B_0).$$

Αν χρησιμοποιήσουμε το θεώρημα της συνολικής μέσης τιμής, έχουμε

$$\begin{aligned} \mathbf{E}[T \mid A_d] = {} & \mathrm{P}(B_d \mid A_d)\mathbf{E}[T \mid A_d \cap B_d] \\ & + \mathrm{P}(B_{d-1} \mid A_d)\mathbf{E}[T \mid A_d \cap B_{d-1}] \\ & + \mathrm{P}(B_{d-2} \mid A_d)\mathbf{E}[T \mid A_d \cap B_{d-2}], \qquad \text{εάν } d > 1, \end{aligned}$$

και

$$\mathbf{E}[T \mid A_1] = \mathrm{P}(B_1 \mid A_1)\mathbf{E}[T \mid A_1 \cap B_1] + \mathrm{P}(B_0 \mid A_1)\mathbf{E}[T \mid A_1 \cap B_0], \quad \text{εάν } d = 1.$$

Με βάση τα δεδομένα του προβλήματος,

$$\mathrm{P}(B_1 \mid A_1) = 2p, \qquad \mathrm{P}(B_0 \mid A_1) = 1 - 2p,$$

$$\mathbf{E}[T \mid A_1 \cap B_1] = 1 + \mathbf{E}[T \mid A_1], \qquad \mathbf{E}[T \mid A_1 \cap B_0] = 1,$$

άρα εφαρμόζοντας τον τύπο για την περίπτωση $d = 1$, έχουμε

$$\mathbf{E}[T \mid A_1] = 2p\big(1 + \mathbf{E}[T \mid A_1]\big) + (1 - 2p),$$

ή

$$\mathbf{E}[T \mid A_1] = \frac{1}{1 - 2p}.$$

Αν εφαρμόσουμε τον τύπο για $d = 2$, έχουμε

$$\mathbf{E}[T \mid A_2] = p\mathbf{E}[T \mid A_2 \cap B_2] + (1 - 2p)\mathbf{E}[T \mid A_2 \cap B_1] + p\mathbf{E}[T \mid A_2 \cap B_0].$$

Έχουμε

$$\begin{aligned} \mathbf{E}[T \mid A_2 \cap B_0] &= 1, \\ \mathbf{E}[T \mid A_2 \cap B_1] &= 1 + \mathbf{E}[T \mid A_1], \\ \mathbf{E}[T \mid A_2 \cap B_2] &= 1 + \mathbf{E}[T \mid A_2], \end{aligned}$$

και αν αντικαταστήσουμε τις σχέσεις αυτές στην έκφραση της $\mathbf{E}[T \mid A_2]$, λαμβάνουμε

$$\begin{aligned}\mathbf{E}[T \mid A_2] &= p\big(1 + \mathbf{E}[T \mid A_2]\big) + (1 - 2p)\big(1 + \mathbf{E}[T \mid A_1]\big) + p \\ &= p\big(1 + \mathbf{E}[T \mid A_2]\big) + (1 - 2p)\left(1 + \frac{1}{1 - 2p}\right) + p.\end{aligned}$$

Η λύση αυτής της εξίσωσης είναι

$$\mathbf{E}[T \mid A_2] = \frac{2}{1 - p}.$$

Αν γενικεύσουμε, έχουμε για $d > 2$,

$$\mathbf{E}[T \mid A_d] = p\big(1 + \mathbf{E}[T \mid A_d]\big) + (1 - 2p)\big(1 + \mathbf{E}[T \mid A_{d-1}]\big) + p\big(1 + \mathbf{E}[T \mid A_{d-2}]\big).$$

Άρα η $\mathbf{E}[T \mid A_d]$ μπορεί να υπολογιστεί αναδρομικά για οποιαδήποτε αρχική απόσταση d, χρησιμοποιώντας σαν αρχικές συνθήκες τις τιμές των $\mathbf{E}[T \mid A_1]$ και $\mathbf{E}[T \mid A_2]$ που βρήκαμε νωρίτερα.

Τέλος, η μέση τιμή της T, χρησιμοποιώντας τη ΣΜΠ για την αρχική απόσταση D και το θεώρημα της συνολικής μέσης τιμής, είναι

$$\mathbf{E}[T] = \sum_d p_D(d)\mathbf{E}[T \mid A_d].$$

Πρόβλημα 35.* Επαληθεύστε τον κανόνα της μέσης τιμής

$$\mathbf{E}\big[g(X, Y)\big] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y),$$

χρησιμοποιώντας τον κανόνα της μέσης τιμής για συναρτήσεις μίας τυχαίας μεταβλητής. Ακολούθως, χρησιμοποιήστε τον κανόνα για την ειδική περίπτωση γραμμικής συνάρτησης, για να επαληθεύσετε τον τύπο

$$\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y],$$

όπου a και b είναι γνωστές σταθερές.

Λύση. Χρησιμοποιούμε το θεώρημα της συνολικής μέσης τιμής και ανάγουμε το πρό-

βλημα στην περίπτωση μίας μοναδικής τυχαίας μεταβλητής. Συγκεκριμένα, έχουμε

$$\begin{aligned}
\mathbf{E}\big[g(X,Y)\big] &= \sum_y p_Y(y)\mathbf{E}\big[g(X,Y)\,|\,Y=y\big] \\
&= \sum_y p_Y(y)\mathbf{E}\big[g(X,y)\,|\,Y=y\big] \\
&= \sum_y p_Y(y)\sum_x g(x,y)p_{X|Y}(x\,|\,y) \\
&= \sum_x \sum_y g(x,y)p_{X,Y}(x,y),
\end{aligned}$$

όπως περιμένουμε. Παρατηρήστε ότι στην τρίτη ισότητα, χρησιμοποιήσαμε τον κανόνα της μέσης τιμής για τη συνάρτηση $g(X,y)$ της μοναδικής τυχαίας μεταβλητής X.

Για την ειδική περίπτωση γραμμικής συνάρτησης , ο κανόνας της μέσης τιμής δίνει

$$\begin{aligned}
\mathbf{E}[aX+bY] &= \sum_x \sum_y (ax+by)p_{X,Y}(x,y) \\
&= a\sum_x x\sum_y p_{X,Y}(x,y) + b\sum_y y\sum_x p_{X,Y}(x,y) \\
&= a\sum_x xp_X(x) + b\sum_y yp_Y(y) \\
&= a\mathbf{E}[X] + b\mathbf{E}[Y].
\end{aligned}$$

Πρόβλημα 36.* Ο κανόνας πολλαπλασιασμού για δεσμευμένες ΣΜΠ. Έστω X, Y και Z τυχαίες μεταβλητές.

(α) Δείξτε ότι

$$p_{X,Y,Z}(x,y,z) = p_X(x)p_{Y|X}(y\,|\,x)p_{Z|X,Y}(z\,|\,x,y).$$

(β) Πως μπορούμε να ερμηνεύσουμε τον τύπο αυτό σαν μία ειδική περίπτωση του κανόνα του πολλαπλασιασμού που δίνεται στην Παράγραφο 1.3;

(γ) Γενικεύστε για την περίπτωση των περισσοτέρων των τριών τυχαίων μεταβλητών.

Λύση. (α) Έχουμε

$$\begin{aligned}
p_{X,Y,Z}(x,y,z) &= \mathrm{P}(X=x, Y=y, Z=z) \\
&= \mathrm{P}(X=x)\mathrm{P}\big(Y=y, Z=z\,|\,X=x\big) \\
&= \mathrm{P}(X=x)\mathrm{P}(Y=y\,|\,X=x)\mathrm{P}(Z=z\,|\,X=x, Y=y) \\
&= p_X(x)p_{Y|X}(y\,|\,x)p_{Z|X,Y}(z\,|\,x,y).
\end{aligned}$$

(β) Ο τύπος μπορεί να γραφτεί ως

$$\mathrm{P}(X=x, Y=y, Z=z) = \mathrm{P}(X=x)\mathrm{P}(Y=y \mid X=x)\mathrm{P}(Z=z \mid X=x, Y=y),$$

που είναι μία ειδική περίπτωση του κανόνα του πολλαπλασιασμού.

(γ) Η γενίκευση είναι

$$\begin{aligned} &p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \\ &\qquad = p_{X_1}(x_1)p_{X_2|X_1}(x_2 \mid x_1)\cdots p_{X_n|X_1,\ldots,X_{n-1}}(x_n \mid x_1,\ldots,x_{n-1}). \end{aligned}$$

Πρόβλημα 37.* Διαίρεση μία τυχαίας μεταβλητής Poisson. Ένας πομπός αποστέλλει είτε ένα 1 με πιθανότητα p, είτε ένα 0 με πιθανότητα $1-p$, ανεξάρτητα από προηγούμενες αποστολές. Εάν ο αριθμός αποστολών μέσα σε ένα δεδομένο χρονικό διάστημα έχει μία Poisson ΣΜΠ με παράμετρο λ, δείξτε ότι ο αριθμός των αποστελλόμενων 1 στο ίδιο χρονικό διάστημα έχει μία Poisson ΣΜΠ με παράμετρο $p\lambda$.

Λύση. Έστω X και Y ο αριθμός των αποστελλόμενων 1 και 0, αντίστοιχα. Έστω $Z = X + Y$ ο συνολικός αριθμός συμβόλων που στάλθηκαν. Έχουμε

$$\begin{aligned} \mathrm{P}(X=n, Y=m) &= \mathrm{P}(X=n, Y=m \mid Z=n+m)\mathrm{P}(Z=n+m) \\ &= \binom{n+m}{n} p^n (1-p)^m \cdot \frac{e^{-\lambda}\lambda^{n+m}}{(n+m)!} \\ &= \frac{e^{-\lambda p}(\lambda p)^n}{n!} \cdot \frac{e^{-\lambda(1-p)}\big(\lambda(1-p)\big)^m}{m!}. \end{aligned}$$

Άρα

$$\begin{aligned} \mathrm{P}(X=n) &= \sum_{m=0}^{\infty} \mathrm{P}(X=n, Y=m) \\ &= \frac{e^{-\lambda p}(\lambda p)^n}{n!} e^{-\lambda(1-p)} \sum_{m=0}^{\infty} \frac{\big(\lambda(1-p)\big)^m}{m!} \\ &= \frac{e^{-\lambda p}(\lambda p)^n}{n!} e^{-\lambda(1-p)} e^{\lambda(1-p)} \\ &= \frac{e^{-\lambda p}(\lambda p)^n}{n!}, \end{aligned}$$

έτσι ώστε η X είναι Poisson με παράμετρο λp.

ΠΑΡΑΓΡΑΦΟΣ 2.7. Ανεξαρτησία

Πρόβλημα 38. Η Αλίκη περνάει τέσσερα φανάρια στο δρόμο για τη δουλειά της. Κάθε φανάρι έχει την ίδια πιθανότητα να είναι πράσινο ή κόκκινο, ανεξάρτητα από τα άλλα.

(α) Ποια είναι η ΣΜΠ, η μέση τιμή και η διασπορά του αριθμού των κόκκινων φαναριών που συναντάει η Αλίκη;

(β) Υποθέστε ότι κάθε κόκκινο φανάρι καθυστερεί την Αλίκη ακριβώς δύο λεπτά. Ποια είναι η διασπορά του χρόνου διαδρομής της Αλίκης;

Πρόβλημα 39. Κάθε πρωί ο Χάρης τρώει αυγά. Σε οποιοδήποτε δεδομένο πρωϊνό, ο αριθμός των αυγών που τρώει έχει την ίδια πιθανότητα να είναι 1, 2, 3, 4, 5, ή 6, ανεξάρτητα από το παρελθόν. Έστω X ο αριθμός των αυγών που ο Χάρης τρώει σε 10 ημέρες. Βρείτε τη μέση τιμή και διασπορά της X.

Πρόβλημα 40. Ένας συγκεκριμένος καθηγητής είναι γνωστός για τον αυθαίρετο τρόπο βαθμολόγησής του. Κάθε εργασία παίρνει ένα βαθμό από το σύνολο $\{A, A-, B+, B, B-, C+\}$, με ίση πιθανότητα, ανεξάρτητα από άλλες εργασίες. Ποια είναι η μέση τιμή του αριθμού εργασιών που θα καταθέσετε πριν πάρετε κάθε δυνατό βαθμό τουλάχιστον μία φορά;

Πρόβλημα 41. Οδηγείτε στη δουλειά σας 5 μέρες την εβδομάδα επί ένα χρόνο (50 εβδομάδες) και με πιθανότητα $p = 0.02$ παίρνετε μία κλήση τροχαίας σε οποιαδήποτε δεδομένη ημέρα, ανεξάρτητα από άλλες ημέρες. Έστω X ο συνολικός αριθμός κλήσεων που παίρνετε σε ένα χρόνο.

(α) Ποια είναι η πιθανότητα ο αριθμός των κλήσεων που παίρνετε να είναι ίσος με τη μέση τιμή της X;

(β) Υπολογίστε προσεγγιστικά την πιθανότητα στο (α) χρησιμοποιώντας την προσέγγιση Poisson.

(γ) Οποιεσδήποτε από τις κλήσεις είναι 10 ή 20 ή 50 ευρώ με αντίστοιχες πιθανότητες 0.5, 0.3 και 0.2 και ανεξάρτητα από άλλες κλήσεις. Να βρείτε τη μέση τιμή και διασπορά του ποσού των χρημάτων που πληρώνετε κατά τη διάρκεια του χρόνου.

(δ) Υποθέστε ότι δεν γνωρίζετε την πιθανότητα p να πάρετε μία κλήση, αλλά πήρατε 5 κλήσεις κατά τη διάρκεια του χρόνου και εκτιμάτε το p με τη δειγματική μέση τιμή

$$\hat{p} = \frac{5}{250} = 0.02.$$

Ποιο είναι το πεδίο δυνατών τιμών της p υποθέτοντας ότι η διαφορά μεταξύ της p και της δειγματικής μέσης τιμής $\hat{p}$ είναι έως 5 φορές την τυπική απόκλιση της δειγματικής μέσης τιμής;

Πρόβλημα 42. Υπολογιστικό πρόβλημα. Θεωρήστε μία πιθανοτική μέθοδο για τον υπολογισμό του εμβαδού ενός δεδομένου υποσυνόλου S του μοναδιαίου τετραγώνου. Η μέθοδος χρησιμοποιεί μία ακολουθία ανεξάρτητων τυχαίων επιλογών σημείων στο μοναδιαίο τετράγωνο $[0, 1] \times [0, 1]$, σύμφωνα με ένα ομοιόμορφο νόμο πιθανότητας. Εάν

το i-οστό σημείο ανήκει στο υποσύνολο S η τιμή της τυχαίας μεταβλητής X_i είναι 1 και διαφορετικά είναι 0. Έστω $X_1, X_2, \ldots$ η ακολουθία τυχαίων μεταβλητών που ορίζεται με τον τρόπο αυτό και για οποιοδήποτε n, έστω

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

(α) Δείξτε ότι η $\mathbf{E}[S_n]$ ισούται με το εμβαδόν του υποσυνόλου S και ότι η $\text{var}(S_n)$ ελαττώνεται στο 0 καθώς το n αυξάνει.

(β) Δείξτε ότι για να υπολογίσετε την τιμή της S_n, αρκεί να γνωρίζετε τις τιμές των S_{n-1} και X_n, ώστε δεν χρειάζεται να θυμάστε τις προηγούμενες τιμές της X_k, $k = 1, \ldots, n-1$. Δώστε ένα τύπο.

(γ) Γράψτε ένα πρόγραμμα που υπολογίζει την τιμή της S_n για $n = 1, 2, \ldots, 10000$, χρησιμοποιώντας τη γεννήτρια τυχαίων αριθμών του υπολογιστή, για την περίπτωση όπου το υποσύνολο S είναι ένας κύκλος εγγεγραμμένος μέσα στο μοναδιαίο τετράγωνο. Πως μπορείτε να χρησιμοποιήσετε το πρόγραμμά σας για να υπολογίσετε πειραματικά την τιμή του π;

(δ) Χρησιμοποιήστε ένα παρόμοιο πρόγραμμα για να υπολογίσετε προσεγγιστικά το εμβαδόν του συνόλου όλων των (x, y) τα οποία βρίσκονται εντός του μοναδιαίου τετραγώνου και ικανοποιούν $0 \leq \cos \pi x + \sin \pi y \leq 1$.

Πρόβλημα 43.* Υποθέστε ότι οι X και Y είναι ανεξάρτητες γεωμετρικές τυχαίες μεταβλητές με την ίδια παράμετρο p. Δείξτε ότι

$$\mathrm{P}(X = i \,|\, X + Y = n) = \frac{1}{n-1}, \qquad i = 1, \ldots, n-1.$$

Λύση. Μπορούμε να ερμηνεύσουμε την $\mathrm{P}(X = i \,|\, X + Y = n)$ ως την πιθανότητα ένα νόμισμα να έρθει κορώνα για πρώτη φορά στην i-οστή ρίψη δεδομένου ότι ήρθε κορώνα για δεύτερη φορά στη n-οστή ρίψη. Ισχυριζόμαστε ότι δεδομένου ότι η δεύτερη κορώνα συνέβη στη n-οστή ρίψη είναι ισοπίθανο η πρώτη κορώνα να έχει έρθει σε οποιαδήποτε ρίψη μεταξύ 1 και $n-1$. Για να το αποδείξουμε αυτό, παρατηρούμε ότι έχουμε

$$\mathrm{P}(X = i \,|\, X + Y = n) = \frac{\mathrm{P}(X = i,\ X + Y = n)}{\mathrm{P}(X + Y = n)} = \frac{\mathrm{P}(X = i)\mathrm{P}(Y = n - i)}{\mathrm{P}(X + Y = n)}.$$

Επίσης

$$\mathrm{P}(X = i) = p(1-p)^{i-1}, \qquad \text{για } i \geq 1,$$

και

$$\mathrm{P}(Y = n - i) = p(1-p)^{n-i-1}, \qquad \text{για } n - i \geq 1.$$

Έπεται ότι

$$\mathrm{P}(X=i)\mathrm{P}(Y=n-i)=\begin{cases} p^2(1-p)^{n-2}, & \text{εάν } i=1,\ldots,n-1,\\ 0, & \text{διαφορετικά.}\end{cases}$$

Επομένως, για οποιαδήποτε i και j στο διάστημα $[1, n-1]$, έχουμε

$$\mathrm{P}(X=i \mid X+Y=n)=\mathrm{P}(X=j \mid X+Y=n),$$

ώστε

$$\mathrm{P}(X=i \mid X+Y=n)=\frac{1}{n-1}, \qquad i=1,\ldots,n-1.$$

Πρόβλημα 44.* Έστω X και Y δύο τυχαίες μεταβλητές με δεδομένη από κοινού ΣΜΠ και έστω g και h δύο συναρτήσεις των X και Y, αντίστοιχα. Δείξτε ότι εάν οι X και Y είναι ανεξάρτητες, τότε το ίδιο ισχύει για τις τυχαίες μεταβλητές $g(X)$ και $h(Y)$.

Λύση. Έστω $U=g(X)$ και $V=h(Y)$. Τότε έχουμε

$$\begin{aligned} p_{U,V}(u,v) &= \sum_{\{(x,y)\,\mid\, g(x)=u,\, h(y)=v\}} p_{X,Y}(x,y)\\ &= \sum_{\{(x,y)\,\mid\, g(x)=u,\, h(y)=v\}} p_X(x)p_Y(y)\\ &= \sum_{\{x\,\mid\, g(x)=u\}} p_X(x) \sum_{\{y\,\mid\, h(y)=v\}} p_Y(y)\\ &= p_U(u)p_V(v), \end{aligned}$$

άρα οι U και V είναι ανεξάρτητες.

Πρόβλημα 45.* Ακραίες διασπορές. Έστω ότι οι $X_1,\ldots,X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές και έστω $X=X_1+\cdots+X_n$ το άθροισμα τους.

(α) Υποθέστε ότι κάθε X_i είναι Bernoulli με παράμετρο p_i και ότι οι $p_1,\ldots,p_n$ επιλέγονται έτσι ώστε η μέση τιμή της X να ισούται με ένα δεδομένο $\mu>0$. Δείξτε ότι η διασπορά της X μεγιστοποιείται εάν οι p_i είναι όλες ίσες με μ/n.

(β) Υποθέστε ότι κάθε X_i είναι γεωμετρική με παράμετρο p_i και ότι οι $p_1,\ldots,p_n$ επιλέγονται έτσι ώστε η μέση τιμή της X να ισούται με ένα δεδομένο $\mu>0$. Δείξτε ότι η διασπορά της X ελαχιστοποιείται εάν οι p_i είναι όλες ίσες με n/μ. [Παρατηρήστε τον πολύ διαφορετικό χαρακτήρα των αποτελεσμάτων (α) και (β).]

Λύση. (α) Έχουμε

$$\mathrm{var}(X)=\sum_{i=1}^{n}\mathrm{var}(X_i)=\sum_{i=1}^{n}p_i(1-p_i)=\mu-\sum_{i=1}^{n}p_i^2.$$

Άρα η μεγιστοποίηση της διασποράς είναι ισοδύναμη με την ελαχιστοποίηση του $\sum_{i=1}^{n} p_i^2$. Μπορεί να επαληθευτεί (χρησιμοποιώντας τον περιορισμό $\sum_{i=1}^{n} p_i = \mu$) ότι

$$\sum_{i=1}^{n} p_i^2 = \sum_{i=1}^{n} (\mu/n)^2 + \sum_{i=1}^{n} (p_i - \mu/n)^2,$$

άρα η $\sum_{i=1}^{n} p_i^2$ ελαχιστοποιείται όταν η $p_i = \mu/n$ για όλα τα i.

(β) Έχουμε

$$\mu = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} \frac{1}{p_i},$$

και

$$\mathrm{var}(X) = \sum_{i=1}^{n} \mathrm{var}(X_i) = \sum_{i=1}^{n} \frac{1 - p_i}{p_i^2}.$$

Αν εισάγουμε την αλλαγή μεταβλητών $y_i = 1/p_i = \mathbf{E}[X_i]$, βλέπουμε ότι ο περιορισμός γίνεται

$$\sum_{i=1}^{n} y_i = \mu,$$

και ότι πρέπει να ελαχιστοποιήσουμε το

$$\sum_{i=1}^{n} y_i(y_i - 1) = \sum_{i=1}^{n} y_i^2 - \mu,$$

υπό τον περιορισμό αυτό. Αυτό είναι το ίδιο με το πρόβλημα του μέρους (α), και εφαρμόζεται η μέθοδος απόδειξης που δίνεται εκεί.

Πρόβλημα 46.* Εντροπία και αβεβαιότητα. Θεωρήστε μία τυχαία μεταβλητή X η οποία μπορεί να πάρει n τιμές, $x_1, \ldots, x_n$, με αντίστοιχες πιθανότητες $p_1, \ldots, p_n$. Η **εντροπία** της X ορίζεται ως

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i.$$

(Όλοι οι λογάριθμοι στο πρόβλημα αυτό θεωρούνται με βάση δύο.) Η εντροπία $H(X)$ παρέχει ένα μέτρο της αβεβαιότητας γύρω από την X. Για να το αντιληφθούμε αυτό, παρατηρήστε ότι $H(X) \geq 0$ και ότι η $H(X)$ είναι πολύ κοντά στο 0 όταν η X είναι "σχεδόν προσδιορισμένη", δηλαδή παίρνει μία από τις δυνατές τιμές της με πιθανότητα πολύ κοντά στο 1 (επειδή έχουμε $p \log p \approx 0$ εάν είτε $p \approx 0$ είτε $p \approx 1$).

Η έννοια της εντροπίας είναι βασική στη θεωρία της πληροφορίας, η οποία άρχισε με την περίφημη δουλειά του C. Shannon και μπορεί να βρεθεί σε πολλά εξειδικευμένα βιβλία. Για παράδειγμα, μπορεί να αποδειχθεί ότι η $H(X)$ είναι ένα κάτω όριο στο μέσο αριθμό των ερωτήσεων με απάντηση ναι-όχι (όπως "είναι η $X = x_1$;" ή "είναι $X < x_5$;") που πρέπει να ερωτηθούν για να προσδιοριστεί η τιμή της X. Επιπλέον, εάν

k είναι ο μέσος αριθμός ερωτήσεων που χρειάζονται για να προσδιοριστεί η τιμή μίας ακολουθίας ανεξάρτητων ίδιας κατανομής τυχαίων μεταβλητών $X_1, X_2, \ldots, X_n$, τότε, με μία κατάλληλη στρατηγική το k/n μπορεί να είναι όσο κοντά στην $H(X)$ επιθυμούμε, όταν το n είναι μεγάλο.

(α) Δείξτε ότι εάν $q_1, \ldots, q_n$ είναι μη αρνητικοί αριθμοί τέτοιοι ώστε $\sum_{i=1}^{n} q_i = 1$, τότε

$$H(X) \leq -\sum_{i=1}^{n} p_i \log q_i,$$

με ισότητα εάν και μόνο εάν $p_i = q_i$ για όλα τα i. Σαν ειδική περίπτωση, δείξτε ότι $H(X) \leq \log n$, με ισότητα εάν και μόνο εάν $p_i = 1/n$ για όλα τα i. *Υπόδειξη*: Χρησιμοποιήστε την ανισότητα $\ln \alpha \leq \alpha - 1$, για $\alpha > 0$, η οποία ισχύει με ισότητα εάν και μόνο εάν $\alpha = 1$.

(β) Έστω X και Y τυχαίες μεταβλητές οι οποίες παίρνουν έναν πεπερασμένο πλήθος τιμών και έχουν από κοινού ΣΜΠ $p_{X,Y}(x, y)$. Ορίστε την

$$I(X, Y) = \sum_x \sum_y p_{X,Y}(x, y) \log \left(\frac{p_{X,Y}(x, y)}{p_X(x) p_Y(y)} \right).$$

Δείξτε ότι $I(X, Y) \geq 0$ και ότι $I(X, Y) = 0$ εάν και μόνο εάν οι X και Y είναι ανεξάρτητες.

(γ) Δείξτε ότι

$$I(X, Y) = H(X) + H(Y) - H(X, Y),$$

όπου

$$H(X, Y) = -\sum_x \sum_y p_{X,Y}(x, y) \log p_{X,Y}(x, y),$$

$$H(X) = -\sum_x p_X(x) \log p_X(x), \qquad H(Y) = -\sum_y p_Y(y) \log p_Y(y).$$

(δ) Δείξτε ότι

$$I(X, Y) = H(X) - H(X \,|\, Y),$$

όπου

$$H(X \,|\, Y) = -\sum_y p_Y(y) \sum_x p_{X|Y}(x \,|\, y) \log p_{X|Y}(x \,|\, y).$$

[Παρατηρήστε ότι η $H(X \,|\, Y)$ μπορεί να θεωρηθεί ως η δεσμευμένη εντροπία της X δεδομένης της Y, δηλαδή, η εντροπία της δεσμευμένης κατανομής της X, δεδομένου ότι $Y = y$, και παίρνοντας το μέσο όρο για όλες τις δυνατές τιμές y. Άρα, η ποσότητα $I(X, Y) = H(X) - H(X \,|\, Y)$ είναι ίση με την ελάττωση της εντροπίας (αβεβαιότητας) της X, όταν η Y γίνεται γνωστή. Επομένως μπορεί να ερμηνευτεί ως η πληροφορία για την X που περιέχεται στην Y και ονομάζεται η **αμφίδρομη**

πληροφορία των X και Y.]

Λύση. (α) Θα χρησιμοποιήσουμε την ανισότητα $\ln\alpha \le \alpha - 1$. (Για να επαληθεύσουμε αυτή την ανισότητα, γράφουμε $\ln\alpha = \int_1^\alpha \beta^{-1}\,d\beta < \int_1^\alpha d\beta = \alpha - 1$ για $\alpha > 1$ και γράφουμε $\ln\alpha = -\int_\alpha^1 \beta^{-1}\,d\beta < -\int_\alpha^1 d\beta = \alpha - 1$ για $0 < \alpha < 1$.)

Έχουμε

$$-\sum_{i=1}^n p_i \ln p_i + \sum_{i=1}^n p_i \ln q_i = \sum_{i=1}^n p_i \ln\left(\frac{q_i}{p_i}\right) \le \sum_{i=1}^n p_i \left(\frac{q_i}{p_i} - 1\right) = 0,$$

με ισότητα εάν και μόνο εάν $p_i = q_i$ για όλα τα i. Εφόσον $\ln p = \log p \ln 2$, παίρνουμε την επιθυμητή σχέση $H(X) \le -\sum_{i=1}^n p_i \log q_i$. Η ανισότητα $H(X) \le \log n$ προκύπτει θέτοντας $q_i = 1/n$ για όλα τα i.

(β) Οι αριθμοί $p_X(x)p_Y(y)$ ικανοποιούν τη σχέση $\sum_x \sum_y p_X(x)p_Y(y) = 1$, άρα από το μέρος (α), έχουμε

$$\sum_x \sum_y p_{X,Y}(x,y) \log\big(p_{X,Y}(x,y)\big) \ge \sum_x \sum_y p_{X,Y}(x,y) \log\big(p_X(x)p_Y(y)\big),$$

με ισότητα εάν και μόνο εάν

$$p_{X,Y}(x,y) = p_X(x)p_Y(y), \qquad \text{για κάθε } x \text{ και } y,$$

το οποίο είναι ισοδύναμο με την ανεξαρτησία των X και Y.

(γ) Έχουμε

$$I(X,Y) = \sum_x \sum_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) - \sum_x \sum_y p_{X,Y}(x,y) \log\big(p_X(x)p_Y(y)\big),$$

και

$$\sum_x \sum_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) = -H(X,Y),$$

$$\begin{aligned}
-\sum_x \sum_y p_{X,Y}(x,y) \log\big(p_X(x)p_Y(y)\big) &= -\sum_x \sum_y p_{X,Y}(x,y) \log p_X(x) \\
&\quad -\sum_x \sum_y p_{X,Y}(x,y) \log p_Y(y) \\
&= -\sum_x p_X(x) \log p_X(x) - \sum_y p_Y(y) \log p_Y(y) \\
&= H(X) + H(Y).
\end{aligned}$$

Αν συνδυάσουμε τις παραπάνω τρεις σχέσεις παίρνουμε $I(X,Y) = H(X) + H(Y) - H(X,Y)$.

(δ) Από τον υπολογισμό του μέρους (γ), έχουμε

$$\begin{aligned}
I(X,Y) &= \sum_x \sum_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) - \sum_x p_X(x) \log p_X(x) \\
&\quad - \sum_x \sum_y p_{X,Y}(x,y) \log p_Y(y) \\
&= H(X) + \sum_x \sum_y p_{X,Y}(x,y) \log \left(\frac{p_{X,Y}(x,y)}{p_Y(y)} \right) \\
&= H(X) + \sum_x \sum_y p_Y(y) p_{X|Y}(x \mid y) \log p_{X|Y}(x \mid y) \\
&= H(X) - H(X \mid Y).
\end{aligned}$$