8.6 Κατανομή Poisson

διακριτές "πολλά μικρά" χρονικά διαστήματα

Είναι προσέγγιση διωνυμικής όταν το $n \to \infty$ και το p (πιθανότητα επιτυχίας) $\to 0$ έτσι ώστε αναλογικά διατηρώντας σε μια σταθερή κατάσταση δηλ. $n \cdot p =$ σταθερό $=$ μέσος όρος

Μετράει τον αριθμό "επιτυχιών" με $n \to \infty$ (σε σχέση με την Binomial)

Τη χρησιμοποιούμε για "σπάνια γεγονότα"
Ουσιαστικά μετράει αριθμό συμβάντων ανά μονάδα χρόνου / εμβαδού / ή ότι άλλο ορίσουμε.

π.χ. α) αριθμός τηλεφωνημάτων σε ένα τηλεφωνικό κέντρο τις 3-5 το απόγευμα.
β) αριθμός αφίξεων στο αερ. Βενιζέλος από τις 7 μέχρι τις 15:00
γ) Ο αριθμός ατυχημάτων από τις 6-9 το πρωί

Ορισμός 8.6.1

Έστω η τυχαία μεταβλητή X με $S_X = \{0, 1, 2, \ldots\}$
Ονομάζεται τυχαία μεταβλητή Poisson με παράμετρο λ αν για κάποιο $\lambda > 0$ ισχύει ότι:

$$P_X(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots$$

πιθανότητα το X να πάρει την τιμή x

Για να εξετάσουμε αν το $P_X(x)$ είναι συνάρτηση πιθανότητας θα πρέπει:
i) Να εξετάσουμε αν το $P_X(x)$ είναι θετικό (όπου σε αυτή την περίπτωση είναι)

ii) Να εξετάσετε αν το $\sum_{x=0}^{\infty} P_X(x)=1$

Πράγματι η

$$\sum_{x=0} P_X(x) = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \overset{e^{\lambda}}{=} e^{-\lambda} e^{\lambda} = 1$$

Άρα η κατανομή Poisson είναι συνάρτηση πιθανότητας.

Άρα συμπεραίνουμε ότι $P(X=i) \approx \frac{\lambda^i e^{-\lambda}}{i!}$

Poisson(λ) Άρα $X \sim Poisson(\lambda)$

ΠΑΡΑΔΕΙΓΜΑΤΑ ΠΟΥ ΑΚΟΛΟΥΘΟΥ ΤΗΝ ΚΑΤΑΝΟΜΗ POISSON

1) Ο αριθμός των τυπογραφικών λαθών ανά σελίδα (ή ανά n σελίδες) σε ένα βιβλίο.

2) Ο αριθμός των ανθρώπων σε μια κοινότητα που η ηλικία τους ξεπερνάει τα 100.

3) Ο αριθμός των πελατών σε ένα ταχυδρομείο ανά ημέρα.

4) Ο αριθμός των α-σωματιδίων που απελευθερώνονται σε ένα συγκεκριμένο χρονικό διάστημα από ένα ραδιενεργό σωματίδιο.

5) Ο αριθμός των τηλεφωνημάτων σε ένα τηλεφωνικό κέντρο ανά ημέρα.

ΜΕΣΗ ΤΙΜΗ ΚΑΙ ΔΙΑΚΥΜΑΝΣΗ POISSON.

$E(X) = \lambda \quad Var(X) = \lambda$

Άρα το λ σημ. τον μέσο όρο των συμβάντων ανά μονάδα μέτρησης.

Εαν δοθεί άσκηση με $n >>>$ και $p <<<$ τότε μπορούμε να χρησιμοποιήσουμε την προσέγγιση Poisson ότι δηλαδή $\lambda = n \cdot p$. Εαν δεν ισχύει $n >>>$ και $p <<<$ τότε ΔΕΝ ΜΠΟΡΩ ΝΑ ΤΟ ΧΡΗΣΙΜΟΠΟΙΗΣΩ.

$[0, t]$ ανά λεπτό με λ μέσο όρο

Τι κατανομή ακολουθεί ο αριθμός αφίξεων στο

Επ. Βενζίνης;

Απ/

Poisson (λt)

α) Μ.Ο 1 άφιξη το λεπτό $\Rightarrow \lambda = 1$

πόσο είναι ανά ώρα $\Rightarrow \lambda' = 60$

β) Μ.Ο 1 άφιξη το 5λεπτό $\Rightarrow \lambda_1 = 5$

ανά ώρα $\Rightarrow \lambda_1' = 12$

ανά 2,5λεπτά $\Rightarrow \lambda_1'' = \frac{1}{2}$

ανά 1 λεπτό $\Rightarrow \lambda_1''' = \frac{1}{5}$

γ) Μ.Ο 3 ατυχήματα την μέρα $\Rightarrow \lambda = 3$

Την εβδομάδα $\Rightarrow \lambda = 3 \cdot 7 = 21$

Άσκ 29 | Παράδειγμα 8.6.1

Έστω ότι στην δυτική Ελλάδα ο αριθμός των σεισμών ακολουθεί την κατανομή Poisson με παράμετρο $\lambda = 2$ ανά εβδομάδα δηλαδή 2 σεισμούς κατά μέσο όρο την εβδομάδα.

Να βρεθεί η πιθανότητα ότι τουλάχιστον 3 σεισμοί θα συμβούν τις επόμενες 2 εβδομάδες.

Λύση

X = ο αριθμός σεισμών ανά εβδομάδα

$X \sim$ Poisson $(\lambda = 2)$

Αν Y = ο αριθμός των σεισμών ανά 2 εβδομάδες τότε $Y \sim$ Poisson $(\lambda' = 2\lambda = 4)$

Άρα $P(Y \geq 3) = 1 - P(Y < 3) =$

$$1 - [P(Y=0) + P(Y=1) + P(Y=2)]$$

$$= 1 - e^{-4} - 4e^{-4} - \frac{16 \cdot e^{-4}}{2} = 1 - e^{-4} - 4e^{-4} - 8e^{-4} =$$

$$= 1 - 13e^{-4} = 0{,}7619.$$

ΚΕΦΑΛΑΙΟ 9
ΑΠΟ ΚΟΙΝΟΥ ΚΑΤΑΝΟΜΗ

Διακριτή τυχαία διμεταβλητή

9.1 Δισδιάστατες διακριτές κατανομές

Έστω X, Y 2 διακριτές τυχαίες μεταβλητές με $S_X = \{0, 1, 2, 3\}$ και $S_Y = \{0, 1\}$

ΠΙΝΑΚΑΣ ⟶ Μας δίνει την από κοινού κατανομή μιας τυχαίας δισδιάστατης μεταβλητής (x, y)

$x \backslash y$	0	1	$P_X(x)$
0	$1/8$	0	$1/8$
1	$2/8$	$1/8$	$3/8$
2	$\frac{1}{8}$	$\frac{2}{8}$	$3/8$
3	0	$1/8$	$1/8$
$P_Y(y)$	$\frac{4}{8}$	$4/8$	1

Κάθε κελί έχει την πιθανότητα εμφάνισης ενός ζεύγους (x, y)

$$P(x,y) = P(X=x, Y=y) = \text{τιμές στον πίνακα}$$

π.χ.

$$P(0,0) = P(X=0, Y=0) = \frac{1}{8}$$

$$P(3,1) = P(X=3, Y=1) = \frac{1}{8}$$

Σε ένα πείραμα μετράμε 2 μεταβλητές όπου το αποτέλεσμα του π.τ. μας δίνεται από ένα διάνυσμα (x, y) μετρήσεων.

Μας αφορούν τυχαίες δισδιάστατες μεταβλητές που ορίζονται στο $\mathbb{R}^2$.

9.2) Ιδιότητες της από κοινού σ.μ.π $p(x,y)$

$S_{xy} = \{(0,0), (0,1), (1,0) \ldots\}$

περιθωριακή σ.μ.π X

i) $p(x,y) \geq 0 \quad \forall (x,y) \in S_{xy}$

ii) $\sum_{\{x,y\} \in S_{xy}} p(x,y) = 1$

9.3) Περιθωριακές σ.μ.π. (συναρτήσεις μάζας πιθανότητας)

i) $p_X(x) = P(X=x) = \sum_{\{y \in S_y\}} p(x,y)$ = άθροισμα στηλών στο πίνακα $p(x,y)$

$P(X=x) = P(X=x, y=0$ ή $X=x, y=1) =$

$= P(X=x, y=0) + P(X=x, y=1) = \sum_{y=0}^{1} p(x,y)$

ii) $p_Y(y)$ = περιθωριακή σ.μ.π της y;

$P(Y=y) = \sum_{x \in S_x} p(x,y)$ = άθροισμα σειρών στον από κοινού πίνακα $p(x,y)$

9.4) Δεσμευμένη πιθανότητα

Έστω ότι γνωρίζουμε ότι $y=0$. Πως αυτό επηρεάζει την πιθανότητα ότι $x=0, 1, 2$ ή 3;

Απάντηση.

Η απάντηση στηρίζεται στον ορισμό της δεσμευμένης πιθανότητας.

Παράδειγμα 9.4.1

$$P\left(X=0/y=0\right)=\frac{P(X=0, y=0)}{P(y=0)}=\frac{P(0,0)}{P(y=0)}=\frac{\frac{1}{8}}{\frac{4}{8}}=\frac{1}{4}$$

Παρατήρηση

$$P\left(X=0/y=0\right)=\frac{1}{4}\neq P(X=0)=\frac{1}{8}$$

Η περιθωριακή πιθανότητα $P(X=0)$ αγνοεί την πληροφορία για την τυχαία μεταβλητή y.

ΟΡΙΣΜΟΣ 9.4.2

α) Η δεσμευμένη σ.μ.π. ισχύει δοθέντος: $Y=y$ $\forall x\in S_x$ δίδεται ως:

$$P(X/y)=P\left[X=x/Y=y\right]=\frac{P(X=x, Y=y)}{P(Y=y)}$$

b) Η δεσμευμένη σ.μ.π. δοθέντος της Y δοθέντος $X=x$ $\forall y\in S_y$ δίδεται ως:

$$P(y/x)=P\left(Y=y/X=x\right)=\frac{P(X=x, Y=y)}{P_X(x)=P(X=x)}$$

ΠΑΝΟΣ
30 Οκτ 2017

9.5 Δεσμευμένη μέση τιμή

$$\mu_{X(x=y)} = E(X / Y=y) = \sum_{\{x \in S_x\}} x \cdot P(X=x / Y=y)$$

$$\mu_Y(y=x) = E(y / X=x) = \sum_{\{y \in S_y\}} y \cdot P(Y=y / X=x)$$

$$E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2}$$

$$E(X / y=0) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot 0 = 1 \; *$$

$$E(X / y=1) = 0 \cdot 0 + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} = 2 \; **$$

$$* \; E\left(\frac{X}{y=0}\right) = 0 \cdot \frac{P(0,0)}{\frac{4}{8}} + 1 \cdot \frac{P(1,0)}{\frac{4}{8}} + 2 \frac{P(2,0)}{\frac{4}{8}} + 3 \frac{P(3,0)}{\frac{4}{8}}$$

$$= 1 \cdot \frac{\frac{2}{\cancel{8}}}{\frac{4}{\cancel{8}}} + 2 \frac{\frac{1}{8}}{\frac{4}{8}} + 3 \frac{0}{\frac{4}{8}} = \frac{1}{2} + \frac{1}{2} = 1$$

$$** \; E\left(\frac{X}{y=1}\right) = 0 \cdot \frac{P(0,1)}{\frac{4}{8}} + 1 \cdot \frac{P(1,1)}{\frac{4}{8}} + 2 \cdot \frac{P(2,1)}{\frac{4}{8}} + 3 \cdot \frac{P(3,1)}{\frac{4}{8}} =$$

$$= 0 + 1 \frac{\frac{1}{\cancel{8}}}{\frac{4}{\cancel{8}}} + 2 \cdot \frac{\frac{2}{\cancel{8}}}{\frac{4}{\cancel{8}}} + 3 \frac{\frac{1}{\cancel{8}}}{\frac{4}{\cancel{8}}} = \frac{1}{4} + 1 + \frac{3}{4} = 2$$

ΠΑΝΟΣ ΣΑΡΑΚΗΝΟΣ

ΑΝΕΞΑΡΤΗΣΙΑ.

$(\forall) = \exists$

Έστω X, Y διακριτές τυχαίες μεταβλητές με σ.μ.π. $p_X(x)$ και $p_Y(y)$, δειγματικούς χώρους S_x, S_y και από κοινού σ.μ.π. $p(x,y)$.

Τότε η X και Y είναι ανεξάρτητες εάν και μόνον εάν $p(x,y) = p_X(x) \cdot p_Y(y)$ $\forall x \in S_x, y \in S_y$

ΑΠΟΤΕΛΕΣΜΑ

Αν X, Y ανεξάρτητες τυχαίες μεταβλητές τότε $p(x/y) = p_X(x)$ $\forall x \in S_x, y \in S_y$

και $p(y/x) = p_Y(y)$ $\forall x \in S_x, y \in S_y$

Διαισθητικά

Γνώση για το X δεν επηρεάζει τις πιθανότητες για το Y.

$p(3,0) = p(x=3, y=0) = 0$

$p(x=3) = \frac{1}{8}$

$p(y=0) = \frac{4}{8}$

$0 \neq \frac{1}{8} \cdot \frac{4}{8}$

9.7 ΣΥΝΔΙΑΚΥΜΑΝΣΗ ΚΑΙ ΣΥΣΧΕΤΙΣΗ

Μέτρηση γραμμικής εξάρτησης 2 τυχαίων μεταβλητών

ΣΥΝΔΙΑΚΥΜΑΝΣΗ (σ_{xy}) Μετρά την κατεύθυνση αλλά όχι την δύναμη της γραμμικής σχέσης 2 τυχαίων μεταβλητών

$$\sigma_{xy} = E\left[(X-\mu_x)\cdot(y-\mu_y)\right] = \sum_{(x,y)\in S_{xy}} (x-\mu_x)(y-\mu_y)\cdot p(x,y)$$

$cov(x,y)$

$$\mu_x = E(x) = \frac{3}{2} \qquad \mu_y = E(y) = \frac{1}{2}$$

$$cov(x,y) = \left(0-\frac{3}{2}\right)\left(0-\frac{1}{2}\right)\cdot\frac{1}{8} + \left(0-\frac{3}{2}\right)\left(1-\frac{1}{2}\right)\cdot 0 +$$

$$+ \dots + \left(3-\frac{3}{2}\right)\left(1-\frac{1}{2}\right)\cdot\frac{1}{8} = \frac{1}{4}$$

ΠΑΝΟΣ ΣΑΡΑΚΗΝΟΣ

ΠΑΝΟΣ
30 Οκτ 2017

ΙΔΙΟΤΗΤΕΣ ΤΗΣ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ

i) $cov(x,y) = cov(y,x)$

ii) $cov(ax, by) = ab \cdot cov(x,y)$ $\quad a, b$ σταθερές

iii) $cov(x,x) = E(x-\mu_x)(x-\mu_x) = E(x-\mu_x)^2 = V(x)$
διακύμανση

iv) x, y ανεξάρτητες $\Rightarrow cov(x,y)=0$

αλλά $cov(x,y)=0 \not\Rightarrow x, y$ ανεξάρτητα

v) ! $\boxed{cov(x,y) = E(xy) - E(x)\cdot E(y)}$

$$E(xy) = \sum_{(x,y)\in \Omega_{xy}} xy \cdot P(x,y) = 0\cdot 0\cdot \frac{1}{8} + 0\cdot 1\cdot 0 + 1\cdot 0\cdot \frac{2}{8} + 1\cdot 1\cdot \frac{1}{8} + 2\cdot 0\cdot \frac{1}{8}$$
$$+ 2\cdot 1\cdot \frac{2}{8} + 3\cdot 0\cdot 0 + 3\cdot 1\cdot \frac{1}{8} =$$
$$= \frac{1}{8} + \frac{4}{8} + \frac{3}{8} = \boxed{1}$$

Άρα $E(xy)=1$ επομένως

$$cov(x,y) = 1 - \frac{3}{2}\cdot\frac{1}{2} = \frac{1}{4}.$$

Αυτό χρησιμεύει καλύτερα για υπολογισμούς
SOS

x , $y=3x$

$$cov(x,y) = cov(x, 3x) = 3cov(x,x) = 3Var(x) = \frac{1}{8} + \frac{9}{8} + \frac{3}{8} = 1$$

ΣΥΣΧΕΤΙΣΗ.

Μετρά την κατεύθυνση όπως και την ισχύ της γραμμικής σχέσης μεταξύ 2 τυχαίων μεταβλητών

$$\rho_{xy} = cor(x,y) = \frac{cov(x,y)}{\sqrt{var(x)}\cdot\sqrt{var(y)}} = \frac{\sigma_{xy}}{\sigma_x\cdot\sigma_y}$$

$$\rho_{xy} = \frac{\frac{1}{4}}{\sqrt{\frac{3}{4}}\cdot\sqrt{\frac{1}{2}}} = 0,577$$

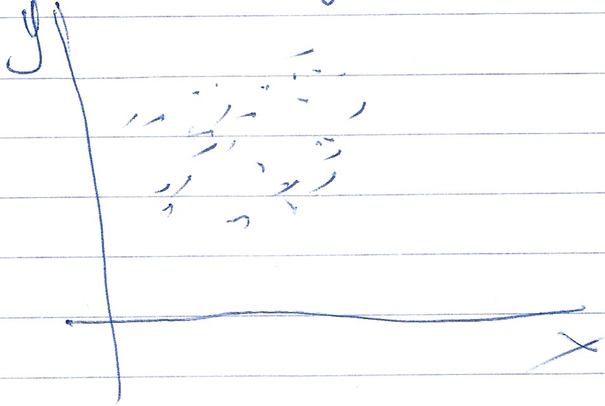
ΙΔΙΟΤΗΤΕΣ ΣΥΣΧΕΤΙΣΗΣ

i) $-1 \le \rho_{xy} \le 1$

ii) $\rho_{xy} = 1$ αν $y = \alpha x + b$ $\alpha > 0$
$\rho_{xy} = -1$ αν $y = \alpha x + b$ $\alpha < 0$

iii) $\rho_{xy} = 0 \Leftrightarrow \sigma_{xy} = 0$

iv) $\rho_{xy} = 0 \not\Rightarrow$ x, y ανεξάρτητα (γενικά)



Θετική συσχέτιση

ΑΣΚΗΣΗ 30 SOS Bernoulli

3 δοκιμές Bernoulli (p)

50% έχουμε $p=\frac{1}{3}$

50% έχουμε $p=\frac{2}{3}$

αβεβαιότητα στο p

Ζητάω την πιθανότητα $P(x=2)$ (να έρθει κ φορές)

Λύση

$$P(x=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}$$

$$P(x=2)=\underbrace{P(x=2/p=\tfrac{1}{3})}_{50\%}+\underbrace{P(x=2/p=\tfrac{2}{3})}_{50\%}$$

Έχω:

~~P(x=2/p=~~

$$P(x=2 \text{ και } p=\tfrac{1}{3})=P\left(\frac{x=2}{p=\frac{1}{3}}\right)\cdot \cancel{P(p=\tfrac{1}{3})}\cdot \underbrace{P(p=\tfrac{1}{3})}_{1/2}=$$

$$=\binom{3}{2}\cdot\left(\frac{1}{3}\right)^2\cdot\left(\frac{2}{3}\right)^1\cdot\frac{1}{2}=\alpha=\frac{2}{9}\cdot\frac{1}{2}=\frac{1}{9}$$

$$P(x=2 \text{ και } p=\tfrac{2}{3})=P\left(\frac{x=2}{p=\frac{2}{3}}\right)\cdot \overbrace{P\left[p=\frac{2}{3}\right]}^{1/2}=$$

$$=\binom{3}{2}\left(\frac{2}{3}\right)^2\left(\cancel{\tfrac{1}{2}}\,\frac{1}{3}\right)^1\cdot\frac{1}{2}=b=\frac{4}{9}\cdot\frac{1}{2}=\frac{2}{9}$$

άρα

$$P(x=2)=\alpha+b \quad (\mathrm{I}) =\frac{1}{9}+\frac{2}{9}=\frac{3}{9}=\frac{1}{3}$$

ΠΡΟΣΟΧΗ. Αν ζητάει $P\left(p=\frac{1}{3}/x=2\right)$ τότε υπολογίζω το (I) και μετά από τύπο BAYES

$$P\left(p=\tfrac{1}{3}/x=2\right)=\frac{P(x=2/p=\frac{1}{3})\cdot P[p=\frac{1}{3}]}{P(x=2)}$$