Στατιστική - Πιθανότητες

Αριθμητικές Μέθοδοι Σύνοψης Δεδομένων

Μη Ομαλοποιημένα Δεδομένα | Ομαλοποιημένα Δεδομένα

Μέτρα θέσης - Κεντρικής Τάσης

Αριθμητικός Μέσος

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\overline{X} \cong \frac{\sum_{i=1}^{k} f_i m_i}{n}, \; n = \sum_{i=1}^{k} f_i$$

Σταθμικός Αριθμητικός Μέσος

$$\overline{X}_w = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}$$

Διάμεσος

Αν n περιττός: M η τιμή της παρατήρησης στη θέση $\frac{n}{2} + \frac{1}{2}$

$$M = X_{\left(\frac{n}{2}+\frac{1}{2}\right)}$$

Αν n άρτιος:

$$M = \frac{1}{2}\left(X_{\left(\frac{n}{2}+1\right)} + X_{\left(\frac{n}{2}\right)}\right)$$

$$M \cong L_M + \frac{\delta}{f_M}\left(\frac{n}{2} - F_{M-1}\right)$$

Επικρατούσα Τιμή

T_0: Η τιμή με τη μεγαλύτερη συχνότητα εμφάνισης

$$T_0 = L_{T_0} + \delta \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

i-Τεταρτημόριο

Το $i = 1, 2, 3$ τεταρτημόριο (Q_i) βρίσκεται στην $[\frac{i(n+1)}{4}]$ θέση. Η τιμή του $i = 1, 2, 3$ τεταρτημορίου (Q_i) είναι

$$Q_i = X_{(A_Q)} + \Delta_Q [X_{(A_Q+1)} - X_{(A_Q)}]$$

όπου A_Q είναι το ακέραιο μέρος του πηλίκου $[\frac{i(n+1)}{4}]$ και Δ_Q είναι το δεκαδικό μέρος του πηλίκου $[\frac{i(n+1)}{4}]$.

$$Q_i = L_{Q_i} + \frac{\delta}{f_{Q_i}}\left(\frac{n \cdot i}{4} - F_{Q_i - 1}\right)$$

Μέτρα Διασποράς

Εύρος

$$R = X_{max} - X_{min}$$

Ενδοτεταρτημοριακό Εύρος

$$IR = Q_3 - Q_1$$

$$IR = Q_3 - Q_1$$

Τεταρτημοριακή Απόκλιση

$$Q = \frac{Q_3 - Q_1}{2}$$

$$Q = \frac{Q_3 - Q_1}{2}$$

Διακύμανση

$$S_{\text{ορ}}^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$$

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

$$S^2 \cong \frac{\sum_{i=1}^{k} f_i(m_i - \overline{X})^2}{n-1}, \ n = \sum_{i=1}^{k} f_i$$

$$S^2 = \frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}{n-1} = \frac{\sum_{i=1}^{n} X_i^2}{n-1} - \frac{(\sum_{i=1}^{n} X_i)^2}{n(n-1)}$$

$$S^2 \cong \frac{\sum_{i=1}^{k} f_i m_i^2 - n\overline{X}^2}{n-1} = \frac{\sum_{i=1}^{k} f_i m_i^2}{n-1} - \frac{(\sum_{i=1}^{k} f_i m_i)}{(n-1)n}$$

Τυπική Απόκλιση

$$S_{\text{ορ}} = +\sqrt{S_{\text{ορ}}^2} \quad \text{ή} \quad S = +\sqrt{S^2}$$

$$S = +\sqrt{S^2}$$

Μέτρα Σχετικής Μεταβλητότητας

Συντελεστής Μεταβλητότητας

$$CV = \frac{S}{X}$$

$$CV = \frac{S}{X}$$

Μέτρα Ασυμμετρίας

Συντελεστές Ασυμμετρίας

$$S_P = \frac{\overline{X} - T_0}{S}$$

$$S_P = \frac{\overline{X} - T_0}{S}$$

$$\beta_3 = \frac{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^3}{n}}{S^3}$$

$$\beta_3 \frac{\frac{\sum_{i=1}^{k} f_i(m_i - \overline{X})^3}{n}}{S^3}$$

Μέτρα Κύρτωσης

Συντελεστής Κύρτωσης

$$\beta_4 = \frac{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^4}{n}}{S^4}$$

$$\beta_4 = \frac{\frac{\sum_{i=1}^{k} f_i(m_i - \overline{X})^4}{n}}{S^4}$$

Πιθανότητες

Έστω S ένας δειγματικός χώρος και έστω $\boldsymbol{B}$ το σύνολο όλων των ενδεχομένων του S. Ορίζουμε ως συνάρτηση πιθανότητας μια συνάρτηση $P\colon \boldsymbol{B} \to \mathbb{R}$, η οποία σε κάθε ενδεχόμενο A αντιστοιχεί έναν πραγματικό αριθμό $P(A)$ έτσι ώστε να ικανοποιούνται τα παρακάτω αξιώματα:

i) $P(A) \geq 0$

ii) $P(S) = 1$

iii) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(A_1) + P(A_2) + \cdots + P(A_n)$

***Θεώρημα* 1.2.1.** *Για κάθε ενδεχόμενο A ισχύει $P(A') = 1 - P(A)$*

***Θεώρημα* 1.2.2.** *Ισχύει ότι $P(\emptyset) = 0$*

***Θεώρημα* 1.2.3.** *Για κάθε ενδεχόμενο A ισχύει $P(A) \leq 1$*

***Θεώρημα* 1.2.4.** *Για οποιαδήποτε ενδεχόμενα A_1 και A_2 ισχύει:*

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Παρατηρήσεις 1.2.1.

- Το θεώρημα 2.2.4 γενικεύεται για την περίπτωση n ενδεχομένων. Στην περίπτωση όπου $n = 3$ γίνεται:

$$\begin{aligned}P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - \\ - P(A_1 \cap A_3) - P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)\end{aligned}$$

- Για δύο ενδεχόμενα τα οποία είναι ασυμβίβαστα το θεώρημα 2.2.4 οδηγεί στο συμπέρασμα:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2), \quad A_1 \cup A_2 = \emptyset$$

  το οποίο είναι ειδική περίπτωση του 3ου αξιώματος.

$$\begin{aligned}P(A_1 \mid A_2) &= \frac{P(A_1 \cap A_2)}{P(A_2)}, \quad P(A_2) \geq 0 \\ P(A_2 \mid A_1) &= \frac{P(A_1 \cap A_2)}{P(A_1)}, \quad P(A_1) \geq 0\end{aligned}$$

- 2 ενδεχόμενα A_1, A_2 είναι *ανεξάρτητα* αν $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$.

- 3 ενδεχόμενα A_1, A_2, A_3 είναι *ανεξάρτητα* αν

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1) \cdot P(A_2) \\ P(A_1 \cap A_3) &= P(A_1) \cdot P(A_3) \\ P(A_2 \cap A_3) &= P(A_2) \cdot P(A_3) \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1) \cdot P(A_2) \cdot P(A_3). \end{aligned}$$

- n ενδεχόμενα $A_1, A_2, \ldots, A_n$ είναι *ανεξάρτητα* αν για κάθε συνδυασμό δύο ή περισσοτέρων από αυτά ισχύει

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k}),$$

με $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n$.

Τα ενδεχόμενα $A_1, A_2, \ldots, A_n$ λέγονται ανεξάρτητα κατά ζεύγη αν

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2), \quad \text{για κάθε} \quad i, j = 1, 2, \ldots, n, \quad i \neq j.$$

Προφανώς, n ενδεχόμενα μπορεί να είναι ανεξάρτητα κατά ζεύγη χωρίς να είναι ανεξάρτητα.

- Για 2 ενδεχόμενα

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 \mid A_1) = P(A_2) \cdot P(A_1 \mid A_2)$$

- Για 3 ενδεχόμενα

$$P[A_1 \cap A_2 \cap A_3] = P[A_1] \cdot P[A_2 \mid A_1] \cdot P[A_3 \mid A_1 \cap A_2]$$

- Για n ενδεχόμενα

$$P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1] \cdot P[A_2 \mid A_1] \cdot P[A_3 \mid A_1 \cap A_2] \cdots P\left[A_n \,\middle|\, \bigcap_{i=1}^{n-1} A_i\right]$$

Έστω ότι $A_1, A_2, \ldots, A_n$ είναι μια διαμέριση του δειγματικού χώρου S τέτοια ώστε $P(A_i) \neq 0$, $i = 1, 2, \ldots, n$. Τότε για κάθε ενδεχόμενο E έχουμε ότι

$$P(E) = \sum_{i=1}^{n} P(A_i) \cdot P(E \mid A_i)$$

Έστω ότι $A_1, A_2, \ldots, A_n$ είναι μια διαμέριση του δειγματικού χώρου S τέτοια ώστε $P(A_i) \neq 0$, $i = 1, 2, \ldots, n$. Τότε για κάθε ενδεχόμενο E με $P(E) > 0$ έχουμε ότι

$$P(A_k \mid E) = \frac{P(A_k) \cdot P(E \mid A_k)}{\sum_{i=1}^{n} P(A_i) \cdot P(E \mid A_i)} = \frac{P(A_k) \cdot P(E \mid A_k)}{P(E)}$$

Αρχές Απαρίθμησης

$$P_n = n!$$

$$\frac{n!}{n_1! \cdot n_2! \cdots n_k!}$$

$$P(n,x) = \frac{n!}{(n-x)!}$$

$$n^x$$

$$C(n,x) = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Κατανομές Πιθανότητας

Η **συνάρτηση πιθανότητας** $P(X = x)$ μιας διακριτής τυχαίας μεταβλητής X ικανοποιεί τις συνθήκες:

i) $P(X = x) \geq 0, \forall x$ στο Πεδίο Ορισμού

ii) $\sum_x P(X = x) = 1$

Η **αθροιστική συνάρτηση κατανομής** $F(a)$ μιας διακριτής τυχαίας μεταβλητής X υπολογίζεται με βάση τη σχέση:

$$F(a) = P(X \leq a) = \sum_{x \leq a} (P(X = x)), \forall a \in \mathbb{R}$$

Η **μέση (αναμενόμενη) τιμή** $\mu = E(X)$ μιας διακριτής τυχαίας μεταβλητής X υπολογίζεται με βάση τον τύπο:

$$E(X) = \sum_x xP(X = x) < +\infty$$

Αν X διακριτή τυχαία μεταβλητή και $g(\cdot)$ μια πραγματική συνάρτηση, τότε η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από τη σχέση:

$$E[g(X)] = \sum_x g(x)P(X = x) < +\infty$$

Η **συνάρτηση πυκνότητας πιθανότητας** $f(x)$ μιας συνεχούς τυχαίας μεταβλητής X ικανοποιεί τις συνθήκες:

i) $f(x) \geq 0, -\infty < x < +\infty$

ii) $\int_{-\infty}^{+\infty} f(x)\, dx = 1$

Η **αθροιστική συνάρτηση κατανομής** $F(a)$ μιας συνεχούς τυχαίας μεταβλητής X υπολογίζεται με βάση τη σχέση:

$$F(a) = P(X \leq a) = \int_{-\infty}^{+\infty} f(t)\, dt, \forall a \in \mathbb{R}$$

Η **μέση (αναμενόμενη) τιμή** $E(X)$ μιας συνεχούς τυχαίας μεταβλητής X υπολογίζεται με βάση τον τύπο:

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\, dx < +\infty$$

Αν X συνεχής τυχαία μεταβλητή και $g(\cdot)$ μια πραγματική συνάρτηση, τότε η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από τη σχέση:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)\, dx < +\infty$$

1. $E[ag(X) + b] = aE[g(X)] + b$, όπου a, b σταθερές.
2. $E[a_1 g_1(x) + a_2 g_2(x)] = a_1 E[g_1(x)] + a_2 E[g_2(x)]$, όπου a_1, a_2 σταθερές.

Έστω X τυχαία μεταβλητή (διακριτή ή συνεχής) με μέση τιμή $\mu = E(X)$. Η **διακύμανση** της X συμβολίζεται με $V(X)$ η σ^2 και δίνεται από τη σχέση:

$$V(X) = \sigma^2 E[X - E(X)]^2 = E(X - \mu)^2 = E(X^2) - \mu^2$$

Ιδιότητα της Διακύμανσης

1. $V(aX + b) = a^2 V(X)$, όπου a, b σταθερές

Όνομα Κατανομής	Παράμετροι	Συνάρτηση πιθανότητας	Σύνολο τιμών	Μέση τιμή	Διασπορά
Διωνυμική $B(n,p)$	$n \in \mathbb{Z}_+$ $0 < p < 1$	$\binom{n}{x} p^x (1-p)^{n-x}$	$x = 0, 1, \ldots, n$	np	$np(1-p)$
Poisson $P(\lambda)$	$\lambda > 0$	$e^{-\lambda} \frac{\lambda^x}{x!}$	$x = 0, 1, \ldots$	λ	λ
Γεωμετρική $G(p)$	$0 < p < 1$	$p(1-p)^{x-1}$	$x = 0, 1, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Υπεργεωμετρική $H(n,a,b)$	$a, b \in \mathbb{Z}_+$ $n \leq a+b$	$\frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}$	$x = 0, 1, \ldots, n$	$\frac{a}{a+b}$	$\frac{nab}{(a+b)^2}$

Όνομα Κατανομής	Παράμετροι	Συνάρτηση πιθανότητας	Σύνολο τιμών	Μέση τιμή	Διασ
Κανονική $N(\mu, \sigma^2)$	$-\infty < \mu < +\infty$ $\sigma > 0$	$\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$-\infty < x < +\infty$	μ	σ
Εκθετική $E(\lambda)$	$\lambda > 0$	$\lambda e^{-\lambda x}$	$0 < x < +\infty$	$\frac{1}{\lambda}$	$\frac{1}{\lambda}$
Ομοιόμορφη $U(a,b)$	$-\infty < a, b < +\infty$	$\frac{1}{b-a}$	$a < x < b$	$\frac{a+b}{2}$	$\frac{(b-}{1}$
Τυποποιημένη Κανονική $N(0,1)$	$-\infty < Z = \frac{X-\mu}{\sigma} < +\infty$	$\frac{1}{\sqrt{2\pi}} e^{\frac{1}{2} z^2}$	$-\infty < z < +\infty$	0	1

Συντελεστές Ευθείας Παλινδρόμησης

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

$$\hat{\beta}_1 = \frac{\sum[(Y_i - \overline{Y})(X_i - \overline{X})]}{\sum[(X_i - \overline{X})^2]} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i^2)}{n}}$$

$$\hat{\beta}_0 = \frac{\sum Y_i}{n} - \hat{\beta}_1 \frac{\sum X_i}{n} = \overline{Y} - \hat{\beta}_1 \overline{X}$$

Συντελεστής Συσχέτισης

$$r = \frac{\sum[(Y_i - \overline{Y})(X_i - \overline{X})]}{\sqrt{\sum(Y_i - \overline{Y})^2 \sum(X_i - \overline{X})^2}} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sqrt{[\sum X_i^2 - \frac{(\sum X_i)^2}{n}][\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}]}}, r \in [-1, 1]$$

Συντελεστής Προσδιορισμού

$$R^2 = \frac{(\sum[(Y_i - \overline{Y})(X_i - \overline{X})])^2}{\sum(Y_i - \overline{Y})^2 \sum(X_i - \overline{X})^2} = \frac{(\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n})^2}{[\sum X_i^2 - \frac{(\sum X_i)^2}{n}][\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}]}, R^2 \in [0, 1]$$

Στατιστική - Πιθανότητες

Αριθμητικές Μέθοδοι Σύνοψης Δεδομένων

Μη Ομαλοποιημένα Δεδομένα | Ομαλοποιημένα Δεδομένα

Μέτρα θέσης - Κεντρικής Τάσης

Αριθμητικός Μέσος

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\overline{X} \cong \frac{\sum_{i=1}^{k} f_i m_i}{n}, \ n = \sum_{i=1}^{k} f_i$$

Σταθμικός Αριθμητικός Μέσος

$$\overline{X}_w = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}$$

Διάμεσος

Αν n περιττός: M η τιμή της παρατήρησης στη θέση $\dfrac{n}{2} + \dfrac{1}{2}$

$$M = X_{\left(\frac{n}{2}+\frac{1}{2}\right)}$$

Αν n άρτιος:

$$M = \frac{1}{2}\left(X_{\left(\frac{n}{2}+1\right)} + X_{\left(\frac{n}{2}\right)}\right)$$

$$M \cong L_M + \frac{\delta}{f_M}\left(\frac{n}{2} - F_{M-1}\right)$$

Επικρατούσα Τιμή

T_0: Η τιμή με τη μεγαλύτερη συχνότητα εμφάνισης

$$T_0 = L_{T_0} + \delta \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

i-Τεταρτημόριο

Το $i = 1, 2, 3$ τεταρτημόριο (Q_i) βρίσκεται στην $[\dfrac{i(n+1)}{4}]$ θέση. Η τιμή του $i = 1, 2, 3$ τεταρτημορίου (Q_i) είναι

$$Q_i = X_{(A_Q)} + \Delta_Q [X_{(A_Q+1)} - X_{(A_Q)}]$$

όπου A_Q είναι το ακέραιο μέρος του πηλίκου $[\dfrac{i(n+1)}{4}]$ και Δ_Q είναι το δεκαδικό μέρος του πηλίκου $[\dfrac{i(n+1)}{4}]$.

$$Q_i = L_{Q_i} + \frac{\delta}{f_{Q_i}}\left(\frac{n \cdot i}{4} - F_{Q_i - 1}\right)$$

Μέτρα Διασποράς

Εύρος

$$R = X_{max} - X_{min}$$

Ενδοτεταρτημοριακό Εύρος

$$IR = Q_3 - Q_1$$

$$IR = Q_3 - Q_1$$

Τεταρτημοριακή Απόκλιση

$$Q = \frac{Q_3 - Q_1}{2}$$

$$Q = \frac{Q_3 - Q_1}{2}$$

Διακύμανση

$$S_{\text{ορ}}^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$$

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

$$S^2 \cong \frac{\sum_{i=1}^{k} f_i(m_i - \overline{X})^2}{n-1},\ n = \sum_{i=1}^{k} f_i$$

$$S^2 = \frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}{n-1} = \frac{\sum_{i=1}^{n} X_i^2}{n-1} - \frac{(\sum_{i=1}^{n} X_i)^2}{n(n-1)}$$

$$S^2 \cong \frac{\sum_{i=1}^{k} f_i m_i^2 - n\overline{X}^2}{n-1} = \frac{\sum_{i=1}^{k} f_i m_i^2}{n-1} - \frac{(\sum_{i=1}^{k} f_i m_i)}{(n-1)n}$$

Τυπική Απόκλιση

$$S_{\text{ορ}} = +\sqrt{S_{\text{ορ}}^2} \quad \text{ή} \quad S = +\sqrt{S^2}$$

$$S = +\sqrt{S^2}$$

Μέτρα Σχετικής Μεταβλητότητας

Συντελεστής Μεταβλητότητας

$$CV = \frac{S}{X}$$

$$CV = \frac{S}{X}$$

Μέτρα Ασυμμετρίας

Συντελεστές Ασυμμετρίας

$$S_P = \frac{\overline{X} - T_0}{S}$$

$$\beta_3 = \frac{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^3}{n}}{S^3}$$

$$S_P = \frac{\overline{X} - T_0}{S}$$

$$\beta_3 \frac{\frac{\sum_{i=1}^{k} f_i(m_i - \overline{X})^3}{n}}{S^3}$$

Μέτρα Κύρτωσης

Συντελεστής Κύρτωσης

$$\beta_4 = \frac{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^4}{n}}{S^4}$$

$$\beta_4 = \frac{\frac{\sum_{i=1}^{k} f_i(m_i - \overline{X})^4}{n}}{S^4}$$

Πιθανότητες

Έστω S ένας δειγματικός χώρος και έστω $\boldsymbol{B}$ το σύνολο όλων των ενδεχομένων του S. Ορίζουμε ως συνάρτηση πιθανότητας μια συνάρτηση $P\colon \boldsymbol{B} \to \mathbb{R}$, η οποία σε κάθε ενδεχόμενο A αντιστοιχεί έναν πραγματικό αριθμό $P(A)$ έτσι ώστε να ικανοποιούνται τα παρακάτω αξιώματα:

i) $P(A) \geq 0$

ii) $P(S) = 1$

iii) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(A_1) + P(A_2) + \cdots + P(A_n)$

Θεώρημα 2.2.1. *Για κάθε ενδεχόμενο A ισχύει $P(A') = 1 - P(A)$*

Θεώρημα 2.2.2. *Ισχύει ότι $P(\emptyset) = 0$*

Θεώρημα 2.2.3. *Για κάθε ενδεχόμενο A ισχύει $P(A) \leq 1$*

Θεώρημα 2.2.4. *Για οποιαδήποτε ενδεχόμενα A_1 και A_2 ισχύει:*

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Παρατηρήσεις 2.2.1.

- Το θεώρημα 2.2.4 γενικεύεται για την περίπτωση n ενδεχομένων. Στην περίπτωση όπου $n = 3$ γίνεται:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - \\ - P(A_1 \cap A_3) - P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

- Για δύο ενδεχόμενα τα οποία είναι ασυμβίβαστα το θεώρημα 2.2.4 οδηγεί στο συμπέρασμα:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2), \quad A_1 \cup A_2 = \emptyset$$

  το οποίο είναι ειδική περίπτωση του 3ου αξιώματος.

$$P(A_1 \mid A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}, \quad P(A_2) \geq 0$$

$$P(A_2 \mid A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}, \quad P(A_1) \geq 0$$

- 2 ενδεχόμενα A_1, A_2 είναι *ανεξάρτητα* αν $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$.

- 3 ενδεχόμενα A_1, A_2, A_3 είναι *ανεξάρτητα* αν

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1) \cdot P(A_2) \\ P(A_1 \cap A_3) &= P(A_1) \cdot P(A_3) \\ P(A_2 \cap A_3) &= P(A_2) \cdot P(A_3) \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1) \cdot P(A_2) \cdot P(A_3). \end{aligned}$$

- n ενδεχόμενα $A_1, A_2, \ldots, A_n$ είναι *ανεξάρτητα* αν για κάθε συνδυασμό δύο ή περισσοτέρων από αυτά ισχύει

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k}),$$

με $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n$.

Τα ενδεχόμενα $A_1, A_2, \ldots, A_n$ λέγονται ανεξάρτητα κατά ζεύγη αν

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2), \quad \text{για κάθε} \quad i, j = 1, 2, \ldots, n, \quad i \neq j.$$

Προφανώς, n ενδεχόμενα μπορεί να είναι ανεξάρτητα κατά ζεύγη χωρίς να είναι ανεξάρτητα.

- Για 2 ενδεχόμενα

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 \mid A_1) = P(A_2) \cdot P(A_1 \mid A_2)$$

- Για 3 ενδεχόμενα

$$P[A_1 \cap A_2 \cap A_3] = P[A_1] \cdot P[A_2 \mid A_1] \cdot P[A_3 \mid A_1 \cap A_2]$$

- Για n ενδεχόμενα

$$P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1] \cdot P[A_2 \mid A_1] \cdot P[A_3 \mid A_1 \cap A_2] \cdots P\left[A_n \,\middle|\, \bigcap_{i=1}^{n-1} A_i\right]$$

Έστω ότι $A_1, A_2, \ldots, A_n$ είναι μια διαμέριση του δειγματικού χώρου S τέτοια ώστε $P(A_i) \neq 0$, $i = 1, 2, \ldots, n$. Τότε για κάθε ενδεχόμενο E έχουμε ότι

$$P(E) = \sum_{i=1}^{n} P(A_i) \cdot P(E \mid A_i)$$

Έστω ότι $A_1, A_2, \ldots, A_n$ είναι μια διαμέριση του δειγματικού χώρου S τέτοια ώστε $P(A_i) \neq 0$, $i = 1, 2, \ldots, n$. Τότε για κάθε ενδεχόμενο E με $P(E) > 0$ έχουμε ότι

$$P(A_k \mid E) = \frac{P(A_k) \cdot P(E \mid A_k)}{\sum_{i=1}^{n} P(A_i) \cdot P(E \mid A_i)} = \frac{P(A_k) \cdot P(E \mid A_k)}{P(E)}$$

Αρχές Απαρίθμησης

$$P_n = n!$$

$$\frac{n!}{n_1! \cdot n_2! \cdots n_k!}$$

$$P(n,x) = \frac{n!}{(n-x)!}$$

$$n^x$$

$$C(n,x) = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Κατανομές Πιθανότητας

Η **συνάρτηση πιθανότητας** $P(X = x)$ μιας διακριτής τυχαίας μεταβλητής X ικανοποιεί τις συνθήκες:

i) $P(X = x) \geq 0, \forall x$ στο Πεδίο Ορισμού

ii) $\sum_x P(X = x) = 1$

Η **αθροιστική συνάρτηση κατανομής** $F(a)$ μιας διακριτής τυχαίας μεταβλητής X υπολογίζεται με βάση τη σχέση:

$$F(a) = P(X \leq a) = \sum_{x \leq a} (P(X = x)), \forall a \in \mathbb{R}$$

Η **μέση (αναμενόμενη) τιμή** $\mu = E(X)$ μιας διακριτής τυχαίας μεταβλητής X υπολογίζεται με βάση τον τύπο:

$$E(X) = \sum_x xP(X = x) < +\infty$$

Αν X διακριτή τυχαία μεταβλητή και $g(\cdot)$ μια πραγματική συνάρτηση, τότε η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από τη σχέση:

$$E[g(X)] = \sum_x g(x)P(X = x) < +\infty$$

Η **συνάρτηση πυκνότητας πιθανότητας** $f(x)$ μιας συνεχούς τυχαίας μεταβλητής X ικανοποιεί τις συνθήκες:

i) $f(x) \geq 0, -\infty < x < +\infty$

ii) $\int_{-\infty}^{+\infty} f(x)\, dx = 1$

Η **αθροιστική συνάρτηση κατανομής** $F(a)$ μιας συνεχούς τυχαίας μεταβλητής X υπολογίζεται με βάση τη σχέση:

$$F(a) = P(X \leq a) = \int_{-\infty}^{+\infty} f(t)\, dt, \forall a \in \mathbb{R}$$

Η **μέση (αναμενόμενη) τιμή** $E(X)$ μιας συνεχούς τυχαίας μεταβλητής X υπολογίζεται με βάση τον τύπο:

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\, dx < +\infty$$

Αν X συνεχής τυχαία μεταβλητή και $g(\cdot)$ μια πραγματική συνάρτηση, τότε η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από τη σχέση:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)\, dx < +\infty$$

1. $E[ag(X) + b] = aE[g(X)] + b$, όπου a, b σταθερές.
2. $E[a_1g_1(x) + a_2g_2(x)] = a_1E[g_1(x)] + a_2E[g_2(x)]$, όπου a_1, a_2 σταθερές.

Έστω X τυχαία μεταβλητή (διακριτή ή συνεχής) με μέση τιμή $\mu = E(X)$. Η **διακύμανση** της X συμβολίζεται με $V(X)$ η σ^2 και δίνεται από τη σχέση:

$$V(X) = \sigma^2 E[X - E(X)]^2 = E(X - \mu)^2 = E(X^2) - \mu^2$$

Ιδιότητα της Διακύμανσης

1. $V(aX + b) = a^2 V(X)$, όπου a, b σταθερές

Όνομα Κατανομής	Παράμετροι	Συνάρτηση πιθανότητας	Σύνολο τιμών	Μέση τιμή	Διασπορά
Διωνυμική $B(n,p)$	$n \in \mathbb{Z}_+$ $0 < p < 1$	$\binom{n}{x} p^x (1-p)^{n-x}$	$x = 0, 1, \ldots, n$	np	$np(1-p)$
Poisson $P(\lambda)$	$\lambda > 0$	$e^{-\lambda} \frac{\lambda^x}{x!}$	$x = 0, 1, \ldots$	λ	λ
Γεωμετρική $G(p)$	$0 < p < 1$	$p(1-p)^{x-1}$	$x = 0, 1, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Υπεργεωμετρική $H(n,a,b)$	$a, b \in \mathbb{Z}_+$ $n \leq a + b$	$\frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}$	$x = 0, 1, \ldots, n$	$\frac{a}{a+b}$	$\frac{nab}{(a+b)^2}$

Όνομα Κατανομής	Παράμετροι	Συνάρτηση πιθανότητας	Σύνολο τιμών	Μέση τιμή	Διασ
Κανονική $N(\mu, \sigma^2)$	$-\infty < \mu < +\infty$ $\sigma > 0$	$\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$-\infty < x < +\infty$	μ	σ
Εκθετική $E(\lambda)$	$\lambda > 0$	$\lambda e^{-\lambda x}$	$0 < x < +\infty$	$\frac{1}{\lambda}$	$\frac{1}{\lambda}$
Ομοιόμορφη $U(a,b)$	$-\infty < a, b < +\infty$	$\frac{1}{b-a}$	$a < x < b$	$\frac{a+b}{2}$	$(b-$ / 1
Τυποποιημένη Κανονική $N(0,1)$	$-\infty < Z = \frac{X-\mu}{\sigma} < +\infty$	$\frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}z^2}$	$-\infty < z < +\infty$	0	1

Συντελεστές Ευθείας Παλινδρόμησης

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

$$\hat{\beta}_1 = \frac{\sum[(Y_i - \overline{Y})(X_i - \overline{X})]}{\sum[(X_i - \overline{X})^2]} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i^2)}{n}}$$

$$\hat{\beta}_0 = \frac{\sum Y_i}{n} - \hat{\beta}_1 \frac{\sum X_i}{n} = \overline{Y} - \hat{\beta}_1 \overline{X}$$

Συντελεστής Συσχέτισης

$$r = \frac{\sum[(Y_i - \overline{Y})(X_i - \overline{X})]}{\sqrt{\sum(Y_i - \overline{Y})^2 \sum(X_i - \overline{X})^2}} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sqrt{[\sum X_i^2 - \frac{(\sum X_i)^2}{n}][\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}]}}, r \in [-1, 1]$$

Συντελεστής Προσδιορισμού

$$R^2 = \frac{(\sum[(Y_i - \overline{Y})(X_i - \overline{X})])^2}{\sum(Y_i - \overline{Y})^2 \sum(X_i - \overline{X})^2} = \frac{(\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n})^2}{[\sum X_i^2 - \frac{(\sum X_i)^2}{n}][\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}]}, R^2 \in [0, 1]$$