# <u>ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ</u>

# <u>ΑΤΑΞΙΝΟΜΗΜΕΝΑ ΔΕΔΟΜΕΝΑ</u>

<u>Αριθμητικός Μέσος:</u> $\frac{\sum Xi}{n}$ όπου n: αριθμός παρατηρήσεων

<u>Διάμεσος:</u> εάν **n** άρτιος $M = \frac{1}{2}\left( X_{\frac{n}{2}+1} + X_{\frac{n}{2}} \right)$

εάν **n** περιττός $M = X_{\frac{n+1}{2}}$

**<u>Παράδειγμα:</u>**

Δηλ.: Εάν n=4 $\rightarrow$ X:1, 2, 3, 4 $\rightarrow$ $M = \frac{1}{2}\left( X_{\frac{n}{2}+1} + X_{\frac{n}{2}} \right)$

$= \frac{1}{2}\left( X_3 + X_2 \right) = \frac{1}{2}(3+2) = \frac{5}{2} = 2{,}5$

Εάν n=3 $\rightarrow$ X:1, 2, 3 $\rightarrow$ $M = X_{\frac{3+1}{2}} = X_{\frac{4}{2}} = X_2 = 2$

<u>Επικρατούσα Τιμή</u>: Η τιμή με την μεγαλύτερη συχνότητα εμφάνισης

<u>Εύρος</u>: Xmax – Xmin

<u>ΤΕΤΑΡΤΗΜΟΡΙΟ – i</u>

$Q_i$ = Η τιμή της παρατήρησης στη θέση $\frac{i(n+1)}{4} \rightarrow \left( X_{\frac{i(n+1)}{4}} \right)$

$Q_i = X_{\frac{i(n+1)}{4}} = X_{AQ} + \Delta_Q (X_{AQ+1} - X_{AQ})$

<u>Όπου:</u> $A_Q$:Το ακέραιο μέρος του πηλίκου $\frac{i(n+1)}{4}$

$\Delta_Q$:Το δεκαδικό μέρος του πηλίκου $\frac{i(n+1)}{4}$

**Παράδειγμα**: $X_i$: 0, 2, 6, 12, 12, 60, 62, 63, 100, 100, 100

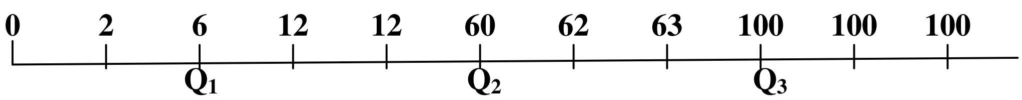
<u>Εύρος</u>: Xmax – Xmin = 100 – 0 =100

<u>Διάμεσος</u>: $X_{\frac{11+1}{2}} = X_{\frac{12}{2}} = X_6 = 60$

<u>Αριθμ. Μέσος</u>: $\frac{\sum Xi}{n} = 47$

$Q_1 = X_{\frac{i(n+1)}{4}} = X_{\frac{1(11+1)}{4}} = X_{\frac{12}{4}} = X_3 \rightarrow 6 \rightarrow Q_1 = 6$

$Q_3 = X_{\frac{i(n+1)}{4}} X_{\frac{3(11+1)}{4}} = X_9 \rightarrow 100 \rightarrow Q_3 = 100$




1° ΤΕΤΑΡΤΗΜΟΡΙΟ: ΤΙΜΗ ΑΡΙΣΤΕΡΑ ΤΗΣ ΟΠΟΙΑΣ ΒΡΙΣΚΕΤΑΙ ΤΟ 25% ΤΟΥ ΣΥΝΟΛΙΚΟΥ ΑΡΙΘΜΟΥ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

2° ΤΕΤΑΡΤΗΜΟΡΙΟ:ΟΡΙΖΕΤΑΙ ΑΠΟ ΤΗΝ ΔΙΑΜΕΣΟ (ΑΡΙΣΤΕΡΑ ΤΟΥ 50%)

3° ΤΕΤΑΡΤΗΜΟΡΙΟ: ΤΙΜΗ ΑΡΙΣΤΕΡΑ ΤΗΣ ΟΠΟΙΑΣ ΒΡΙΣΚΕΤΑΙ ΤΟ 75% ΤΩΝ ΔΕΔΟΜΕΝΩΝ

<u>Ενδοτεταρτημοριακό Εύρος:</u> IR = $Q_3$ - $Q_1$ = 100 - 6 = 94 ΕΙΝΑΙ ΣΧΕΔΟΝ ΙΣΟ ΜΕ ΤΟ ΕΥΡΟΣ

<u>Διακύμανση</u>: $S^2 = \dfrac{\sum X_i^2 - n\overline{X}}{n-1}$ <u>Τυπική Απόκλιση:</u> $S = \sqrt{S^2}$

**ΠΑΡΑΔΕΙΓΜΑ**:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $X_i$: | 0 | 2 | 6 | 12 | 12 | 60 | 62 | 63 | 100 | 100 | 100 | $\Sigma X_i = 517$ |
| $X_i^2$: | 0 | 4 | 36 | 144 | 144 | 3600 | 3844 | 3969 | 10000 | 10000 | 10000 | $\Sigma X_i^2 = 41741$ |

$\overline{X} = \dfrac{\sum X_i}{n} = \dfrac{517}{11} = 47$ $S^2 = \dfrac{\sum X_i^2 - n\overline{X}}{n-1} = \dfrac{41741 - 11 * 47}{10} = 4122{,}9$ $S = \sqrt{S^2} = \sqrt{4122{,}4} = 64{,}21$

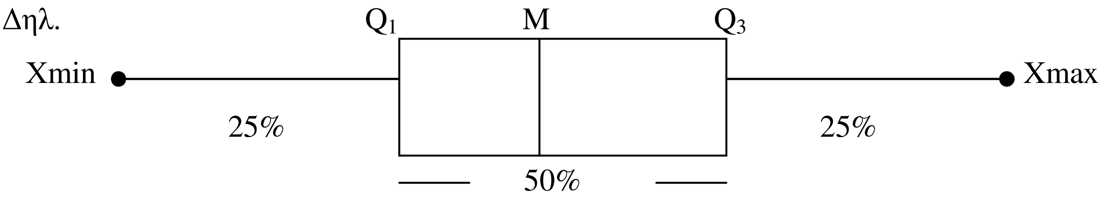
<u>Συντελεστής Ασυμμετρίας</u>: $Sp = \dfrac{\overline{X} - T_o}{S}$ ή $Sp = \dfrac{3(\overline{x} - M)}{S}$

όπου $\overline{X} \rightarrow$ Μέσος, $T_0 \rightarrow$ Επικρατούσα Τιμή, S $\rightarrow$ Τυπική Απόκλιση

Άρα: $Sp = \dfrac{\overline{X} - T_0}{S} = \dfrac{47 - 100}{64{,}21} = -0{,}825$: Η κατανομή των δεδομένων εμφανίζει αρνητική ασυμμετρία

<u>ΘΗΚΟΓΡΑΜΜΑ</u>: Το θηκόγραμμα αποτελεί γραφικό τρόπο παρουσίασης 5 περιληπτικών μέτρων μιας κατανομής: Xmax**,** Xmin, M, $Q_3$, $Q_1$

Δηλ.



# ΤΑΞΙΝΟΜΗΜΕΝΑ ΔΕΔΟΜΕΝΑ

| ΤΑΞΗ | ΚΕΝΤΡΙΚΗ ΤΙΜΗ ΤΑΞΗΣ $m_i$ | ΣΥΧΝΟΤΗΤΑ fi | $m_i \times f_i$ | ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ $F_i$ | $m_i$ - 16,23 | $(m_i - 16{,}23)^2$ | fi $(m_i – 16{,}23)^2$ |
|---|---|---|---|---|---|---|---|
| 10-12 | 11 | 4 | 44 | 4 | -5.23 | 27.35 | 109.4 |
| 12-14 | 13 | 21 | 273 | 25 | -3.23 | 10.43 | 219.03 |
| 14-16 | 15 | 57 | 855 | 82 | -1.23 | 1.51 | 86.07 |
| 16-18 | 17 | 39 | 663 | 121 | 0.77 | 0.59 | 23.01 |
| 18-20 | 19 | 22 | 418 | 143 | 2.77 | 7.67 | 168.74 |
| 20-22 | 21 | 8 | 168 | 151 | 4.77 | 22.75 | 182 |
| 22-24 | 23 | 2 | 46 | 153 | 6.77 | 45.83 | 91.66 |
| 24-26 | 25 | 2 | 50 | 155 | 8.77 | 76.91 | 153.82 |
| ΣΥΝΟΛΑ | | 155 | 2517 | | | | 1033.73 |

Αριθμητικός Μέσος: $\overline{X} = \dfrac{\sum f_i m_i}{\sum f_i} = \dfrac{2517}{155} = 16.23$

Επικρατούσα Τιμή: $T_0 = L_{T_0} + \delta \dfrac{\Delta_1}{\Delta_1 + \Delta_2}$

* Η τάξη που περιέχει την επικρατούσα τιμή είναι αυτή με την μεγαλύτερη συχνότητα

$L_{T_0}$ : Κατώτερο άκρο της τάσης που περιέχει την επικρατούσα τιμή

$\Delta_1$: Η διαφορά συχνοτήτων της ομάδας που περιέχει την επικρατούσα τιμή και της προηγούμενης ομάδας

$\Delta_2$: Η διαφορά των συχνοτήτων μεταξύ της ομάδας που περιέχει την επικρατούσα τιμή και της επόμενης ομάδας

$T_0 = L_{T_0} + \delta \dfrac{\Delta_1}{\Delta_1 + \Delta_2} = 14 + 2 \dfrac{57-21}{(57-21)+(57-39)} = 15.33$

$L_{T_0}$ =14 διότι η τάξη με την μεγαλύτερη συχνότητα είναι η (14-16)

δ=2, $\Delta_1$=57-21, $\Delta_2$=57-39

Διάμεσος: $M = L_M + \delta \frac{(\frac{n}{2} - F_{M-1})}{f_M}$

*Η τάξη που περιέχει τον διάμεσο εντοπίζεται σαν αυτή που έχει αθροιστική συχνότητα $\geq \frac{n}{2}$

$L_M$: Κατώτερο άκρο της τάξης που περιέχει το διάμεσο (Μ)

$F_{M-1}$: Αθροιστική συχνότητα πριν από την ομάδα που περιέχει τον διάμεσο

$f_M$: Συχνότητα της ομάδας που περιέχει τον διάμεσο

$M = L_M + \delta \frac{(\frac{n}{2} - F_{M-1})}{f_M} = 14 + 2 \frac{(\frac{155}{2} - 25)}{57} = 15.84$

$\frac{n}{2} = \frac{155}{2} = 77{,}5$ άρα η τάξη (14-16) έχει αθροιστική συχνότητα = 82 > 77,5

$L_M = 14$, $F_{M-1} = 25$, $f_M = 57$

ΤΕΤΑΡΤΗΜΟΡΙΟ – i

**$Q_1$**

$Q_1 = L_{Q1} + \delta \frac{(\frac{n}{4} - F_{Q_{1-1}})}{f_{Q_1}}$

*Η τάξη που περιέχει το $Q_1$ εντοπίζεται σαν αυτή που έχει αθροιστική συχνότητα $\geq \frac{n}{4}$

Δηλ. $\frac{155}{4} = 38.75$ και είναι η (14-16) $\leftrightarrow$ $F_{Q1} = 82$

$L_{Q1}$: Κατώτερο άκρο της τάξης που περιέχει το $Q_1$ ($L_{Q1}=14$)

$F_{Q1-1}$: Αθροιστική συχνότητα πριν από την ομάδα που περιέχει το $Q_1$ ($F_{Q_{1-1}}$ =25)

$f_{Q1}$: Συχνότητα της ομάδας που περιέχει το $Q_1$ ($f_{Q1}=57$)

Άρα με: $L_{Q1}=14$, δ=2, $F_{Q_{1-1}}$ =25, n=155 → $Q_1 = 14 + 2 \frac{(\frac{155}{4} - 25)}{57} = 14{,}48$

**<u>$Q_3$</u>**

$Q_3 = L_{Q3} + \delta \frac{(\frac{3n}{4} - F_{Q3-1})}{f_{Q3}}$

*Η τάξη που περιέχει το $Q_3$ εντοπίζεται σαν αυτή που έχει αθροιστική συχνότητα $\geq \frac{3n}{4}$

Δηλ. $\frac{3*155}{4} = 116,25$ και είναι η (16-18) $\leftrightarrow$ $F_{Q3} = 121$

$L_{Q3}$: Κατώτερο άκρο της τάξης που περιέχει το $Q_3$ ($L_{Q3}=16$)

$F_{Q3-1}$: Αθροιστική συχνότητα πριν από την ομάδα που περιέχει το $Q_3$ ($F_{Q3-1}=82$)

$f_{Q3}$: Συχνότητα της ομάδας που περιέχει το $Q_3$ ($f_{Q3}=39$)

Άρα με: $L_{Q3}=16$, $F_{Q3-1}= 82$, $f_{Q3}=39$ → $Q_3 = L_{Q3} + \delta \frac{(\frac{3n}{4} - F_{Q3-1})}{f_{Q3}} = 16+2\frac{\frac{3*155}{4} - 82}{39} = .......$

<u>Διακύμανση</u>: $S^2 = \frac{\sum f_i (m_i - \overline{X})^2}{\sum f_i - 1} = \frac{\sum f_i (m_i - 16.23)^2}{\sum f_i - 1} = \frac{1033.73}{155 - 1} = 6.71$

<u>Τυπική Απόκλιση</u>: $S = \sqrt{S^2} = \sqrt{6.71} = 2.59$

<u>Συντελεστής Μεταβλητότητας</u>: $CV = \frac{S}{\overline{X}} * 100$

**Παράδειγμα**:

Δεύτερο δείγμα με: $\overline{X}=16$ και S=3 → Να συγκριθεί με το προηγούμενο: $\overline{X}=16.23$ και S=2.59

$CV_A = \frac{2.59}{16.23} * 100 = 15.958\%$ & $CV_B = \frac{3}{16} * 100 = 18.75\%$

Το δεύτερο δείγμα έχει μεγαλύτερη διασπορά εν σχέσει με τον πρώτο

# <u>ΠΙΘΑΝΟΤΗΤΕΣ</u>

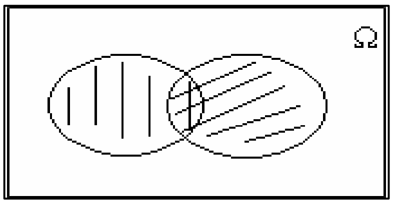
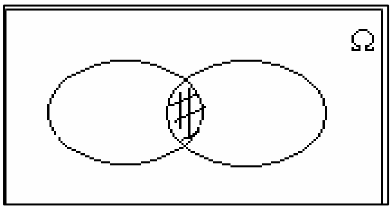
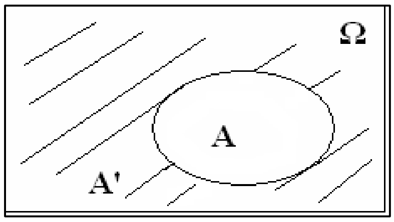
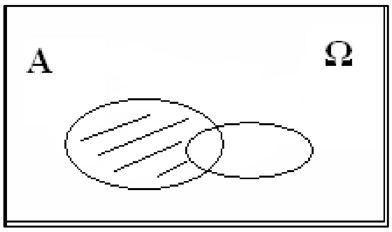
<u>Δειγματικός Χώρος</u>: Το σύνολο των δυνατών αποτελεσμάτων ενός τυχαίου πειράματος αποτελεί το δειγματικό χώρο Ω.

<u>Δειγματικό Σημείο ή Στοιχειώδες Ενδεχόμενο</u>: Κάθε ένα από τα δυνατά αποτελέσματα ενός τυχαίου πειράματος ονομάζεται στοιχειώδες ενδεχόμενο ή δειγματικό σημείο.

<u>Ενδεχόμενο</u>: Κάθε υποσύνολο Α του δειγματικού χώρου Ω ονομάζεται ενδεχόμενο.

**<u>Παράδειγμα</u>**: Η ρίψη ενός νομίσματος 2 φορές: $\Omega = \{ΚΓ, ΓΓ, ΚΚ, ΓΚ\}$

Ενδεχόμενο: $A = \{ΚΚ, ΓΓ\} \subseteq \Omega$

| Συμβολισμός | Ενδεχόμενο | Σημασία | Παράσταση |
|---|---|---|---|
| $A \cup B$ | " Α ή Β " | Το $A \cup B$ πραγματοποιείται όταν πραγματοποιείται ένα τουλάχιστον από τα Α,Β |   |
| $A \cap B$ | " Α και Β " | Το $A \cap B$ πραγματοποιείται όταν πραγματοποιούνται συγρόνως το Α και Β |   |
| Α' | Όχι Α ή αντίθετο του Α ή συμπληρωματικό του Α | Το Α΄ πραγματοποιείται όταν πραγματοποιείται το Α |   |
| Α-Β | Η διαφορά του Β από το Α | Το Α-Β πραγματοποιείται, όταν πραγματοποιείται το Α αλλά όχι το Β |   |

Δύο ενδεχόμενα Α και Β είναι ασυμβίβαστα όταν $A \cap B = \varnothing$

Έστω Ω = $\{w_1, w_2, \ldots\ldots.., w_\nu\}$ ένας δειγματικός χώρος με πεπερασμένο πλήθος στοιχείων. Σε κάθε απλό ενδεχόμενο αντιστοιχεί μια πιθανότητα του ενδεχομένου $\{w_i\}$. Επίσης ισχύει $0 \leq P(w_i) \leq 1$ και $P(w_1) + P(w_2) + P(w_3) + \ldots + P(w_\nu) = 1$.

<u>ΠΡΟΣΘΕΣΗ ΠΙΘΑΝΟΤΗΤΩΝ</u>

2 ΕΝΔΕΧΟΜΕΝΑ:

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

3 ΕΝΔΕΧΟΜΕΝΑ:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)$$

Για 2 ενδεχόμενα Α, Β για τα οποία $A \subseteq B$ ισχύει: $P(A) \leq P(B)$

Για κάθε ενδεχόμενο Α ισχύει: $P(A) + P(A') = 1$

ΔΕΣΜΕΥΜΕΝΗ ΠΙΘΑΝΟΤΗΤΑ: Αν Α και Β είναι δύο ενδεχόμενα ενός δειγματικού χώρου Ω και P(B) > 0, τότε ο λόγος $\frac{P(A\cap B)}{P(B)}$ λέγεται δεσμευμένη πιθανότητα του Α με δεδομένο το Β και συμβολίζεται με P(A| B). Δηλαδή: P(A| B)= $\frac{P(A\cap B)}{P(B)}$ με P(B)>0.

Άρα $P(A\cap B$ =P(A|B)*P(B)= P(B| A) * P(A) Πολλαπλασιαστικός Νόμος των Πιθανοτήτων

**ΠΑΡΑΔΕΙΓΜΑ**: Μια κάλπη περιέχει 7 σφαίρες μπλε και 3 κόκκινες. Δύο σφαίρες εκλέγονται τυχαία, διαδοχικά χωρίς επανατοποθέτηση. Θέλουμε να υπολογίσουμε την πιθανότητα η πρώτη σφαίρα να είναι κόκκινη και η δεύτερη μπλε.

ΛΥΣΗ: Έστω: Α – Η σφαίρα είναι κόκκινη και Β – Η σφαίρα είναι μπλε. Η πιθανότητα η πρώτη σφαίρα να είναι κόκκινη είναι P(A) = $\frac{3}{10}$. Η δεύτερη σφαίρα επιλέγεται από την κάλπη που περιέχει 7 μπλε σφαίρες και 2 κόκκινες. Η πιθανότητα η δεύτερη σφαίρα να είναι μπλε είναι: P(B|A)=7/9. Άρα η πιθανότητα η πρώτη επιλογή να είναι κόκκινη σφαίρα και η δεύτερη μπλε είναι: P(A$\cap$B)=P(A)*P(B|A)=(3/10)*(7/9)=7/30

ΑΝΕΞΑΡΤΗΤΑ ΕΝΔΕΧΟΜΕΝΑ: P(A| B) = P(A) ή P(B| A) = P(B)

$$P(A\cap B) = P(A) * P(B)$$

$$P(A\cap B\cap \Gamma) = P(A) * P(B) * P(\Gamma)$$

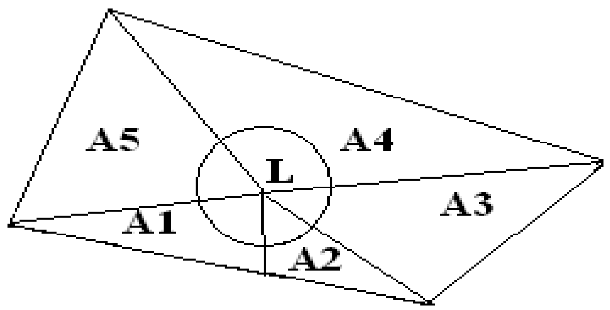
**Παράδειγμα**: Έστω ότι βγάζουμε ένα χαρτί από μια τράπουλα και θέλουμε το χαρτί αυτό να είναι άσσος σπαθί. Ποια η πιθανότητα;

Λύση: A = {Το χαρτί είναι άσσος} $\leftrightarrow P(A) = \frac{4}{52} = \frac{1}{13}$

B = {Το χαρτί είναι σπαθί} $\leftrightarrow P(B) = \frac{13}{52} = \frac{1}{4}$

$P(A\cap B)$ = P(A) * P(B) = $\frac{1}{13} * \frac{1}{4} = \frac{1}{52}$

ΘΕΩΡΗΜΑ ΟΛΙΚΗΣ ΠΙΘΑΝΟΤΗΤΑΣ: Έστω ότι $A_1$, $A_2$, ……… $A_n$ είναι μία διαμέριση του δειγματικού χώρου S τέτοια ώστε P(Ai) ≠ 0, i=1,2,…….,n τότε για κάθε ενδεχόμενο Ε έχουμε: P(E) = $\sum_{i=1}^{n} P(A_i) * P(E/A_i)$




$P(E) = P(A_1) * P(E|A_1) + P(A_2) * P(E|A_2) + P(A_3) * P(E|A_3) + P(A_4) * P(E|A_4) + P(A_5) * P(E|A_5)$

**Παράδειγμα**: Τρία κουτιά περιέχουν στοιχεία, μερικά από τα οποία είναι ελαττωματικά. Η αναλογία φαίνεται στον παρακάτω πίνακα.

| | ΑΡΙΘΜΟΣ ΣΤΟΙΧΕΙΩΝ | ΑΡΙΘΜΟΣ ΕΛΑΤΤΩΜΑΤΙΚΩΝ |
|---|---|---|
| ΚΟΥΤΙ 1 | 10 | 4 |
| ΚΟΥΤΙ 2 | 6 | 1 |
| ΚΟΥΤΙ 3 | 8 | 3 |

Διαλέγουμε ένα κουτί στην τύχη και στην συνέχεια διαλέγουμε ένα στοιχείο στην τύχη από το κουτί αυτό. Να βρεθεί η πιθανότητα το στοιχείο να είναι ελαττωματικό.

Λύση: Έστω Ai = {το στοιχείο προέρχεται από το κουτί i} i = 1,2,3

E = {το στοιχείο να είναι ελαττωματικό}

$P(E) = P(E|A_1) * P(A_1) + P(E|A_2) * P(A_2) + P(E|A_3) * P(A_3) = (4 / 10) * (1/3) + (1/6) * (1/3) + (3/8) * (1/3) = 113 / 360$

ΘΕΩΡΗΜΑ BAYES: Έστω $A_1$, $A_2$, ……… $A_n$ μια διαμέριση του δειγματικού χώρου S με $P(A_i) > 0$ για κάθε i=1,2,….,n τότε, για κάθε ενδεχόμενο Ε με P(E) > 0 έχουμε ότι:

$$P(A_\kappa|E) = \frac{P(E/A_k)*P(A_k)}{\sum_i P(E/A_i)*P(A_i)} = \frac{P(E/A_k)*P(A_k)}{P(E)}$$

**Παράδειγμα**: Στην αρχή του χρόνου, διατυπώθηκαν τρεις οικονομικές θεωρίες για την πιθανή εξέλιξη της Ελληνικής οικονομίας. Όταν διατυπώθηκαν και οι τρεις θεωρίες φαίνονταν ισοπιθανές. Στο τέλος του έτους εξετάσθηκε η πραγματική κατάσταση της οικονομίας με αναφορά τις τρεις θεωρίες. Η ανάλυση κατέληξε στο συμπέρασμα ότι αν η πρώτη θεωρία ήταν αληθινή η οικονομία θα είχε πιθανότητα 0,6 να καταλήξει στην παρούσα κατάσταση. Οι αντίστοιχες πιθανότητες για την δεύτερη και την τρίτη πρόβλεψη είναι 0,4 και 0,2. Να υπολογισθεί η πιθανότητα με την οποία η παρούσα κατάσταση της οικονομίας μπορεί να θεωρηθεί αποτέλεσμα της θεωρίας i, i = 1,2,3

Λύση: Έστω Ai = {η θεωρία i είναι σωστή} i = 1,2,3

Έχουμε: $P(A_1) = P(A_2) = P(A_3) = 1/3$

Έστω Ε ={Η οικονομία βρίσκεται στην παρούσα κατάσταση}

Με βάση το θεώρημα Bayes έχουμε για την πρώτη θεωρία:

$$P(A_1/E) = \frac{P(E/A_1)*P(A_1)}{P(E/A_1)*P(A_1)+P(E/A_2)*P(A_2)+P(E/A_3)*P(A_3)} = \frac{\left(\frac{6}{10}\right)*\left(\frac{1}{3}\right)}{\left(\frac{6}{10}\right)*\left(\frac{1}{3}\right)+\left(\frac{4}{10}\right)*\left(\frac{1}{3}\right)+\left(\frac{2}{10}\right)*\left(\frac{1}{3}\right)}$$

$$= \frac{6}{12} = \frac{1}{2}$$

κλπ. για τις άλλες θεωρίες

# **<u>ΚΑΤΑΝΟΜΕΣ</u>**

**<u>Διωνυμική Κατανομή</u>**: $P_{n,x}$ = P(X=x) = $\binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$

<u>Μέση Τιμή:</u> μ = np

<u>Τυπική Απόκλιση:</u> σ = $\sqrt{npq}$

**<u>ΠΑΡΑΔΕΙΓΜΑ</u>**: Ένα εργοστάσιο παράγει προϊόντα από τα οποία το 10% είναι ελαττωματικά. Έστω ότι επιλέγονται τυχαία 4 προϊόντα. Ποια η πιθανότητα από το δείγμα των 4 συγκεκριμένων προϊόντων τα δυο να είναι ελαττωματικά;

<u>ΛΥΣΗ:</u> x=2, n=4, p=0.10, q=0.90

P(x=2) = $\binom{n}{x} p^x q^{n-x} \Rightarrow P(x=2) = \binom{4}{2} * 0{,}10^2 * 0{,}90^{4-2} = 0{,}049 \Rightarrow P(x=2) = 4{,}860\%$

Όταν Ερωτηθείτε: Ποια η πιθανότητα από το δείγμα των 4 προϊόντων τουλάχιστον δύο να είναι ελαττωματικά. Τότε πρέπει να βρείτε τις εξής πιθανότητες: P(x=2), P(x=3), P(x=4) και άρα η P(x≥2)=P(x=2)+P(x=3)+P(x=4). Ή να βρείτε P(x=0), P(x=1) και άρα η P(x≥2) = 1 – P(x=0) – P(x=1)

**<u>Κατανομή Poisson</u>:** $P(x=\kappa) = \dfrac{e^{-\lambda}\lambda^{\kappa}}{\kappa!}$ κ= 0,1,2……

Μέση Τιμή: μ=λ

Τυπική Απόκλιση: σ=λ

**<u>Παράδειγμα</u>**: Μεταξύ των ωρών 6μμ και 7μμ η υπηρεσία καταλόγου Αττικής του ΟΤΕ δέχεται κατά μέσο όρο <u>2 κλήσεις το λεπτό</u>. Υποθέτοντας ότι οι κλήσεις κατανέμονται τυχαία στο χρόνο, βρείτε την πιθανότητα η τηλεφωνήτρια της συγκεκριμένης υπηρεσίας να δεχθεί <u>σε κάποιο τυχαία επιλεγμένο λεπτό:</u> α) 4 κλήσεις, β)6 κλήσεις σε τυχαία περίοδο δύο λεπτών.

<u>Λύση:</u> α)Συμβολίζουμε με x τον αριθμό των κλήσεων που γίνονται σε τυχαίο λεπτό

$$P[x=4] = \frac{e^{-2} * 2^{4}}{4!} = 0{,}090$$

$$\beta)\ P[x=6] = \frac{e^{-2} * 2^{6}}{6!} = 0{,}104$$

**<u>Υπεργεωμετρική Κατανομή:</u>** P(x=κ) = $\dfrac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \dfrac{C(r,k)*C(N-r,n-k)}{c(N,n)} = \dfrac{C_k^r * C_{n-k}^{N-r}}{C_n^N}$

Όπου: $\binom{r}{k} = C(r,k) = C_k^r = \dfrac{r!}{k!(r-k)!}$

**<u>Παράδειγμα</u>**: Σε μια στέρνα υπάρχουν N=15 ψάρια από τα οποία r=5 είναι κόκκινα και τα υπόλοιπα N-r=10 είναι μαύρα. Ποια είναι η πιθανότητα στα n=4 ψάρια που θα πιάσουμε τα 3 να είναι κόκκινα;

<u>Λύση</u>: κ=3, N=15, r=5, N-r=10, n=4

P(x=3) = $\dfrac{\binom{5}{3}\binom{15-5}{4-3}}{\binom{15}{4}} = \dfrac{\binom{5}{3}\binom{10}{1}}{\binom{15}{4}} = 0{,}033$

Η πιθανότητα στα 4 ψάρια που θα πιάσουμε τα 3 να είναι κόκκινα είναι 0,033

# <u>ΣΥΝΔΥΑΣΜΟΙ</u>

<u>Ορισμός</u>: Έστω ένα σύνολο S περιέχει n στοιχεία. Ένα υποσύνολο του S από r στοιχεία είναι ένας συνδυασμός των n ανά r. Το σύνολο των διαφορετικών συνδυασμών των n ανά r συμβολίζεται με: $\binom{n}{r}$ ή $C_r^n$ ή C (r, n)

<u>Πόρισμα</u>: $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

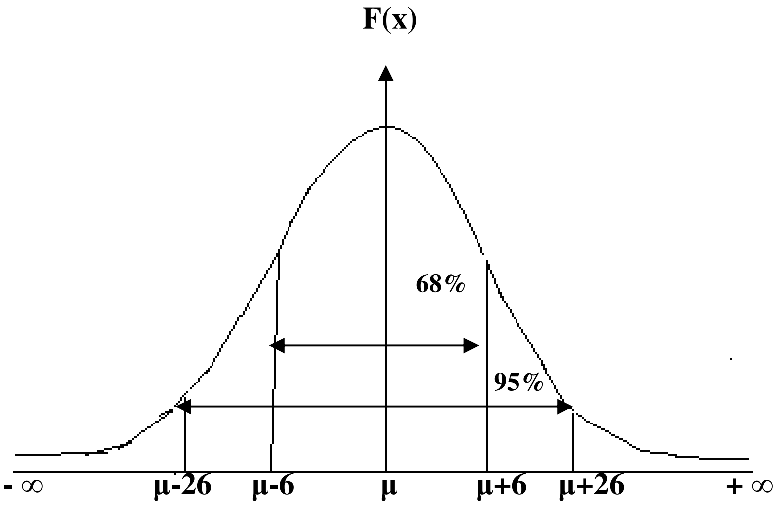
**<u>Παράδειγμα</u>**: Να βρεθεί ο αριθμός των τριμελών επιτροπών που είναι δυνατόν να ορισθούν από ένα σύνολο 8 ατόμων

<u>Λύση</u>: $\binom{8}{3} = \frac{8!}{3!(8-3)!} = 56$

**<u>Παράδειγμα:</u>** Ένας μαθητής πρέπει να απαντήσει στις εξετάσεις της ιστορίας σε 6 από 9 ερωτήσεις. Πόσες επιλογές έχει;

<u>Λύση</u>: $\binom{9}{6} = \frac{9!}{6!(9-6)!} = \frac{9!}{6!3!} = \frac{6!\cdot 7\cdot 8\cdot 9}{6!3!} = \frac{7\cdot 8\cdot 9}{1\cdot 2\cdot 3} = 84$

# ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ




- Η κανονική κατανομή είναι συμμετρική περί το μ που είναι ο μέσος όρος όλων των δυνατών τιμών της τ.μ.χ.
- Η επικρατούσα τιμή, η διάμεσος και ο μέσος ταυτίζονται λόγω συμμετρίας της καμπύλης της κατανομής
- Το συνολικό εμβαδόν που περικλείεται από την καμπύλη και τον άξονα των Χ είναι μοναδιαίο

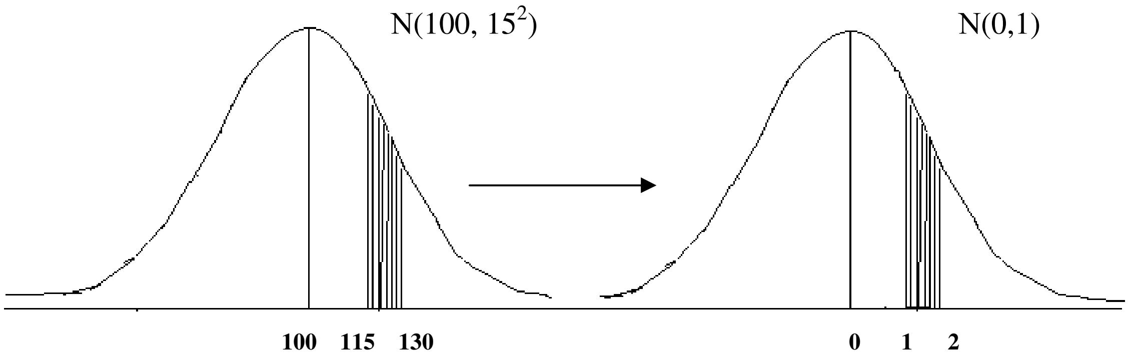
ΕΠΙΣΗΣ: $X \sim N(\mu,\sigma^2)$, $E[X]=\mu$, $Var(X)=\sigma^2$

**<u>Τυποποιημένη Κανονική Κατανομή</u>**: Αν η τυχαία μεταβλητή X έχει μέσο μ και διακύμανση $\sigma^2$, τότε η τυχαία μεταβλητή $Z = \dfrac{X-\mu}{\sigma}$ αποκλείει την τυποποιημένη μορφή της X και έχει μέσο ίσο με μηδέν (0) και διακύμανση ίση με τη μονάδα (1) $\rightarrow Z$ ~N(0,1)

* Αν η τυχαία μεταβλητή X~N(μ, $\sigma^2$), τότε η τυχαία μεταβλητή Z~N(0,1)

**<u>Παράδειγμα</u>**: Η τυχαία μεταβλητή X~N(100, $15^2$), να υπολογισθεί η πιθανότητα P(115 < X < 130)

<u>Απάντηση</u>: $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-100}{15}$ ~ N(0,1)

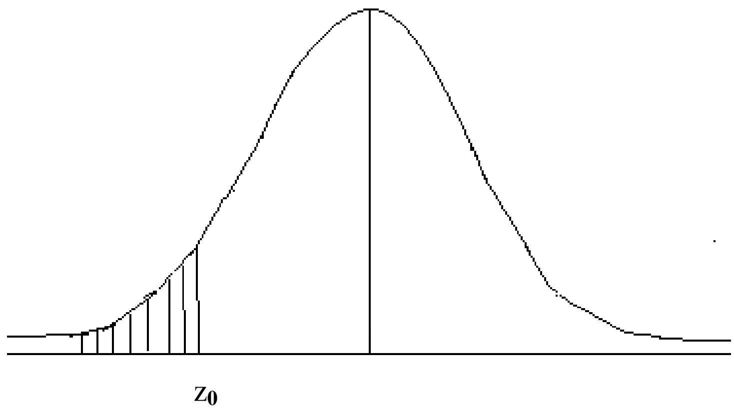



P(115<X<130) = P $\left(\dfrac{115-100}{15} < \dfrac{X-100}{15} < \dfrac{130-100}{15}\right)$ = P(1<Z<2)

Το σκιαγραφημένο εμβαδόν της N(100, $15^2$) ισοδυναμεί με το σκιαγραφημένο κόκκινο εμβαδόν της N(0,1). <u>Πρέπει πρώτα να μιλήσουμε για την αθροιστική κατανομή προκειμένου να υπολογίσουμε την πιο πάνω πιθανότητα.</u> Είναι ανάγκη λοιπόν να ορίσουμε τη συνάρτηση <u>αθροιστικής κατανομής</u> ή απλώς <u>συνάρτηση κατανομής</u> της τυποποιημένης κανονικής κατανομής N(0,1).

**Αθροιστική Κατανομή ή Συνάρτηση Κατανομής της Τυποποιημένης Κανονικής Κατανομής N(0,1).**

Η συνάρτηση κατανομής της N(0,1) συμβολίζεται με Φ(z) και ορίζεται από την $P(-\infty<Z<z) = P(Z<z)$. Δηλαδή, για κάποιο $z_0$ η $P(-\infty<Z<z_0) = \Phi(z_0)$

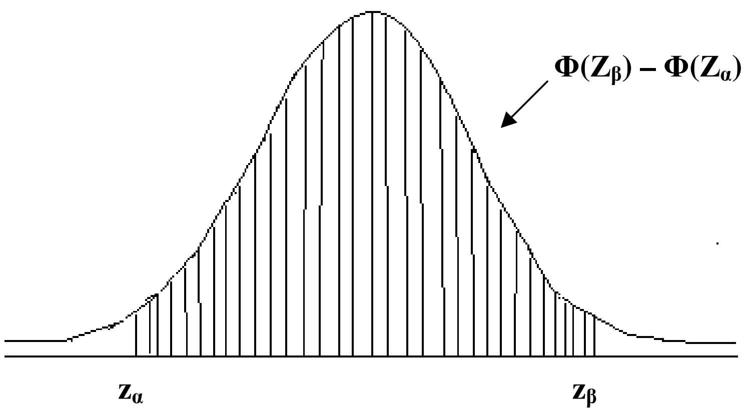



Βασιζόμενοι λοιπόν στον ορισμό της Φ(z) είναι φανερό ότι οι πιθανότητες της μορφής $P(\alpha<X<\beta)$, μετασχηματιζόμενες σε πιθανότητες της μορφής $P(z_\alpha<Z< z_\beta)$ υπολογίζονται ως ακολούθως:

$$P(\alpha<X<\beta) = P\left(\frac{\alpha-\mu}{\sigma}\langle\frac{X-\mu}{\sigma}\langle\frac{\beta-\mu}{\sigma}\right) = P(z_\alpha<Z<z_\beta)$$

$$P(-\infty<Z<z_\beta) - P(-\infty<Z<z_\alpha) = \Phi(z_\beta) - \Phi(z_\alpha)$$

Όπου $z_\alpha = \frac{\alpha-\mu}{\sigma}$ και $z_\beta = \frac{\beta-\mu}{\sigma}$ είναι οι τυποποιημένες τιμές της μεταβλητής $X \sim N(\mu, \sigma^2)$



<u>Λύση του Προηγούμενου Παραδείγματος</u>:

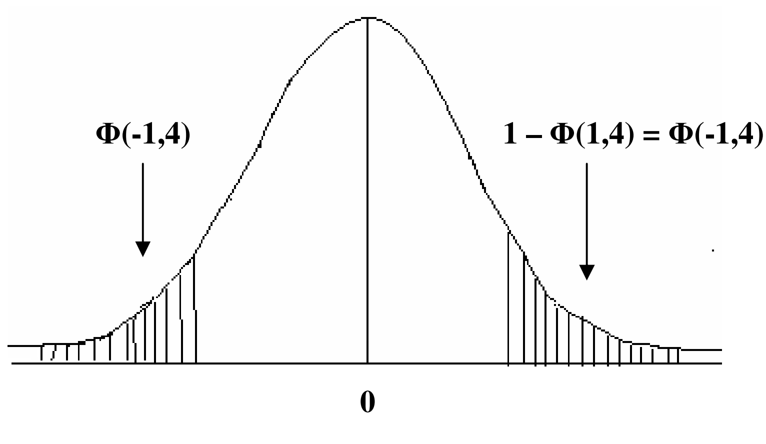
$P(115<X<130) = P\left(\frac{115-100}{15} \langle \frac{X-100}{15} \langle \frac{130-100}{15}\right) = P(1<Z<2) = P(-\infty<Z<2) - P(-\infty<Z<1) = \Phi(2) - \Phi(1) = 0{,}9772 - 0{,}8413 = 0{,}135$

Η εύρεση τιμών της Φ(Ζ) διευκολύνεται από την ύπαρξη πινάκων. Από τους πίνακες Φ(2) = 0,9772 και Φ(Α) = 0,8413.

Επίσης: **Φ(-Z) = 1-Φ(Z)**

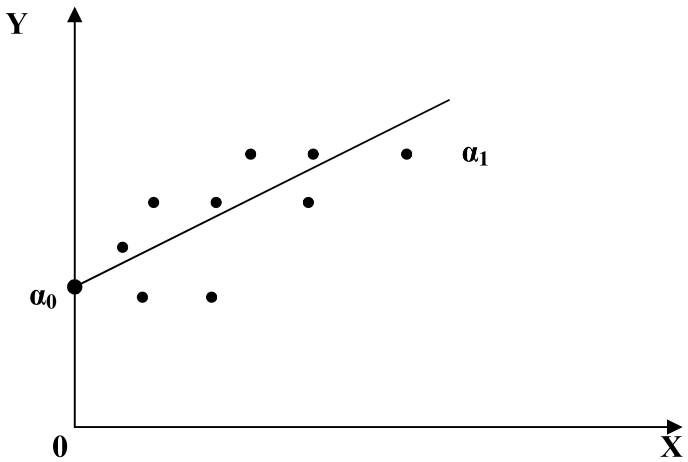
**<u>Παράδειγμα:</u>** Να βρεθεί η πιθανότητα $P(-\infty<Z<-1{,}4)$

<u>Λύση:</u> $P(-\infty<Z<-1{,}4) = \Phi(-1{,}4) = 1 - \Phi(1{,}4) = 1 - 0{,}9192 = 0{,}0808$



# ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

| ΜΗΝΕΣ | ΠΩΛΗΣΕΙΣ | ΔΙΑΦΗΜΙΣΗ | | | |
|---|---|---|---|---|---|
| | Y | X | $X^2$ | $Y^2$ | XY |
| 1 | 48.0 | 3.2 | 10.24 | 2304 | 153.6 |
| 2 | 74.5 | 4.1 | 16.81 | 5478 | 305.45 |
| 3 | 35.7 | 1.1 | 1.21 | 1274.49 | 39.27 |
| 4 | 33.4 | 2.0 | 4.0 | 1115.57 | 66.8 |
| 5 | 60.0 | 3.3 | 10.89 | 3600 | 198 |
| 6 | 77.3 | 5.6 | 31.36 | 5975.29 | 432.88 |
| 7 | 55.8 | 2.5 | 6.25 | 3113.64 | 139.5 |
| 8 | 58.2 | 3.8 | 14.44 | 3387.24 | 221.16 |
| 9 | 60.1 | 4.8 | 23.04 | 3612.01 | 288.48 |
| 10 | 92.4 | 4.8 | 23.04 | 8537.76 | 443.52 |
| 11 | 63.7 | 3.9 | 15.21 | 4057.69 | 248.43 |
| 12 | 44.0 | 2.2 | 4.84 | 1936 | 96.8 |
| ΑΘΡΟΙΣΜΑ | 703.1 | 30.9 | 161.33 | 44389.68 | 2633.89 |




$$\hat{Y} = \alpha_0 + \alpha_1 X$$

$$a_1 = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}, \quad a_0 = \frac{\sum Y}{n} - a_1 \frac{\sum X}{n}$$

Άρα : $a_1 = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sum X^2 - \frac{\left(\sum X\right)^2}{n}} = \frac{2633.89 - \frac{(703.1)(30.9)}{12}}{(161.33) - \frac{(30.9)^2}{12}} = \frac{2633.89 - 1810.48}{161.33 - 79.57} \Rightarrow$

$$\alpha_1 = \frac{823.41}{81.76} \Rightarrow \alpha_1 = 10$$

$$\alpha_0 = \frac{703.1}{12} - 10 * \frac{30.9}{12} = 58.59 - 10 * 2.575 \Rightarrow a_0 = 32.84$$

$\hat{Y} = \alpha_0 + \alpha_1 X \Rightarrow \hat{Y} = 32.84 + 10 * X$

Χρησιμοποιώντας το τυπολόγιο βρίσκουμε τον **Συντελεστή Συσχέτισης** $\rightarrow r$ και τον **Συντελεστή Προσδιορισμού** $R^2 = r^2$

Εναλλακτικά για την εκτίμηση του $α_1$ μπορούμε να χρησιμοποιήσουμε τον τύπο των αποκλίσεων των τιμών από τους μέσους:

$$\alpha_1 = \frac{\sum\left[\left(\mathrm{Y}_i - \overline{Y}\right)\left(X_i - \overline{X}\right)\right]}{\sum\left[\left(X_i - \overline{X}\right)^2\right]}$$

| X | Y | $X - \overline{Y}$ | $Y - \overline{Y}$ | $\left(X - \overline{X}\right)^2$ | $\left(Y - \overline{Y}\right)^2$ | $\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)$ |
|---|---|---|---|---|---|---|
| 3 | 10 | -0.75 | -0.75 | 0.5625 | 0.5625 | 0.5625 |
| 4 | 12 | 0.25 | 1.25 | 0.0625 | 1.5625 | 0.3125 |
| 6 | 17 | 2.25 | 6.25 | 5.0625 | 39.0625 | 14.0625 |
| 2 | 4 | -1.75 | -6.76 | 3.0625 | 45.5625 | 11.8125 |
| | | | | 8.75 | 86.75 | 26.75 |

$$\overline{X} = \frac{\sum X}{n} = \frac{15}{4} = 3.75$$

$$\overline{Y} = \frac{\sum Y}{n} = \frac{43}{4} = 10.75$$

$$\alpha_1 = \frac{\sum\left[\left(\mathrm{Y}_i - \overline{Y}\right)\left(X_i - \overline{X}\right)\right]}{\sum\left[\left(X_i - \overline{X}\right)^2\right]} = \frac{26.75}{8.75} \Rightarrow a_1 = 30.57$$

**Επίσης βασιζόμενοι στο τυπολόγιο μπορούμε να υπολογίσουμε r και $R^2$ χρησιμοποιώντας τους τύπους των αποκλίσεων των τιμών από τους μέσους.**