Estadísitca i Probabilitat

μφ COMBINATORIA

Conceptualització

L'Estadística és l'estudi de la recollida, organització, anàlisi i interpretació de dades. És l'àrea matemàtica més interdisciplinar, essent usada extensivament no només en Física sino en totes les ciències naturals i socials.

$$\text{Variable Estadística}\begin{cases}\text{Qualitativa}\\ \text{Quantitativa}\begin{cases}\text{Discreta}\\ \text{Contínua}\end{cases}\end{cases}$$

Dades No Agrupades

Per un conjunt de dades $\{x_1, x_2, \dots, x_n\} \in \mathbb{R}$, definim:

Mitjana Aritmètica

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Desviació Mitjana respecte la Mitjana

$$D_{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$$

Variança

$$\mathcal{V} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

Desviació Típica o Estàndar

$$\sigma = \sqrt{\mathcal{V}} = \left(\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2\right)^{1/2}$$

Dades Agrupades

Per un conjunt de dades numèriques que estiguin agrupades en una taula de freqüències amb c classes i amb freqüències absolutes respectives n_i, $\{\underbrace{x_1, \dots, x_1}_{n_1}, \underbrace{x_2, \dots, x_2}_{n_2}, \dots, \underbrace{x_c, \dots, x_c}_{n_c}\} \in \mathbb{R}$, definim:

Mitjana Aritmètica

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{c} x_i \cdot n_i$$

Desviació Mitjana respecte la Mitjana

$$D_{\bar{x}} = \frac{1}{n}\sum_{i=1}^{c} |x_i - \bar{x}| \cdot n_i$$

Variança

$$\mathcal{V} = \frac{1}{n}\sum_{i=1}^{c} (x_i - \bar{x})^2 \cdot n_i$$

Estadística Bidimensional

Similarment al cas unidimensional, si tenim dues variables estadistiques x i y relacionades entre si, es poden calcular tots els parametres estadístics per separat per a cadascuna d'aquestes variables: $\bar{x}$, $\bar{y}$, σ_x i σ_y. A partir d'aquestes últimes es poden construir dos paràmetres estadístics més:

Covariança

$$\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{c} (x_i - \bar{x}) \cdot (y_i - \bar{y}) \cdot n_i = \left[\frac{1}{n}\sum_{i=1}^{n} x_i \cdot y_i\right] - \bar{x} \cdot \bar{y}$$

Coeficient de Correlació

$$r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{\displaystyle\sum_{i=1}^{c} (x_i - \bar{x}) \cdot (y_i - \bar{y}) \cdot n_i}{\left[\displaystyle\sum_{i=1}^{c} (x_i - \bar{x})^2 \cdot n_i\right]^{1/2} \cdot \left[\displaystyle\sum_{i=1}^{c} (y_i - \bar{y})^2 \cdot n_i\right]^{1/2}}$$

El coeficient de correlació pot prendre qualsevol valor dins de l'interval $[-1, 1]$. Prendrà els valors extrems -1 i 1 quan hi hagi una màxima correlació (les dues variables depenen fortament l'una de l'altre) positiva o negativa (dependència decreixent o creixent). Prendrà el valor de 0 quan no hi hagi cap relació entre els valors d'una variable i l'altre, quan siguin completament independents.

Recta de Regressió

Quan dues variables estadístiques tinguin una certa correlació, es podrà establir una llei lineal entre elles a través de la construcció d'una recta de regressió que lligui les dues variables:

$$y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Longleftrightarrow y = \bar{y} + \frac{\sigma_y}{\sigma_x} r (x - \bar{x})$$

Estadísitca i Probabilitat

μφ COMBINATORIA

Conceptualització

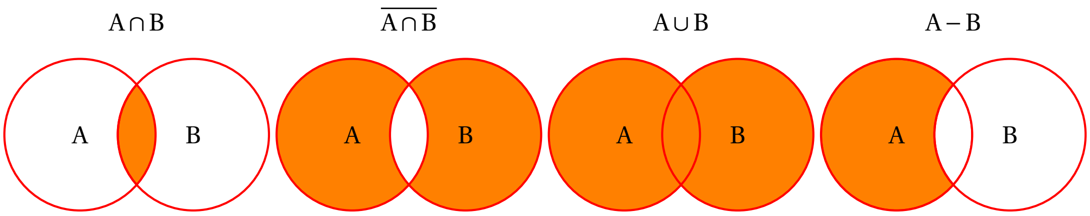
La Teoria de la Probabilitat estudia les relacions lògiques i aritmètiques entre successos, que ocorren amb certeses quantificables però no absolutes, en el marc d'un **Espai de Probabilitat** $(\Omega, \mathfrak{E}, \mathfrak{P})$ que consta de:

- **Events**: conjunts de resultats que poden succeir en l'execució d'un cert experiment.
- un **Espai de Successos** $\mathfrak{E}$: un conjunt d'events que poden contenir qualsevol nombre de resultats possibles.
- un **Espai Mostral** Ω: conjunt de tots els resultats possibles.
- una **Mesura de Probabilitat** $\mathfrak{P}$: funció que assigna probabilitats a cadascun del resultats possibles.

Els events es poden composar entre si amb els operadors lògics habituals de la Teoria de Conjunts:

- $A \cap B$: la **Intersecció** entre A i B és el conjunt de successos que pertanyen simultàniament als dos events
- $A \cup B$: la **Unió** dels events A i B és el conjunt de tots els successos d'un i altre events
- $A - B$: la **Substracció** d'un event B de A és el conjunt de successos d'A excepte els que pertanyen també a B.
- $\bar{A}$: la **Negació** d'un cert event A és el conjunt de tots els successos que no pertanyin a A

Aquestes operacions entre events es poden visualitzar fàcilment a través dels **Diagrames de Venn** habituals:



Axiomes de Kolmogorov

- Primer Axioma: $P(E) \in \mathbb{R} \wedge P(E) \geq 0 \quad \forall E \in \mathfrak{E}$
  La probabilitat de qualsevol event E que pertanyi a $\mathfrak{E}$ ha de ser un nombre finit positiu o zero
- Segon Axioma: $P(\Omega) = 1$
  La probabilitat total ha de ser normalitzada a 1. $P(\Omega)$ es pot pensar com la probabilitat d'un succés segur, doncs sempre es donarà algun resultat o altre de l'espai mostral
- Tercer Axioma: $P(E_1 \cup E_2 \cup \cdots) = \sum_{i=1}^{\infty} P(E_i)$
  Per a una seqüència contable no necessàriament finita d'events mutuament excloents, la probabilitat de la unió de tots els successos esdevé la suma de les probabilitats de cada event per separat

A partir dels tres axiomes se'n dedueixen altres propietats respecte a la probabilitat de diferents events:

$$A \subseteq B \implies P(A) \leq P(B) \quad \natural \quad P(\emptyset) = 0 \quad \natural \quad 0 \leq P(E) \leq 1 \quad \natural \quad P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \natural \quad P(\Omega \backslash E) = 1 - P(E)$$

Probabilitat Condicionada i Diagrames d'Arbre

La probabilitat condicionada fa referència a com la probabilitat que es dongui un succés concret A pot veure's alterada si en un event anterior s'ha donat el succés B. Aquesta llei ens permet calcular les probabilitats de composició d'events sabent les probabilitats condicionades de cada event a través de simples diagrames d'arbre:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \implies \boxed{P(A \cap B) = P(B) \cdot P(A/B)}$$