TEMA 7: Estadística i probabilitat

ESTADISTICA

7.1 CONCEPTES BÀSICS

La **Estadística** tracta del record, ordenació y classificació de les dades obtingudes per les observacions, per poder fer comparacions i treure'n conclusions

- Un **estudi estadístic** consta de les següents fases:

  Recollida de dades.

  Organització y representació de dades.

  Anàlisis de dades.

  Obtenció de conclusions.

- **Població**

Una **població** és el conjunt de tots els elements objecte de l'estudi estadístic

- **Individu**

Un **individu** o **unitat estadística** és cadascun dels elements que componen la població .

- **Mostra**

Una **mostra** es un conjunt representatiu de la població de referència, el nombre de individus de una mostra es menor que el de la població.

- **Mostreig**

El **mostreig** és la reunió de dades que es desitja estudiar, obtinguts d'una porció reduïda i representativa de la població.

- **Valor**

Un **valor** és cadascun dels diferents resultats que s'obtenen en un estudi estadístic. Si llancem una moneda a l'aire 5 cops podem obtenir dos valors: cara i creu.

- **Dada**

Una **dada** és cadascun dels valors que se han obtingut al realitzar un estudi estadístic. Si llancem una moneda a l'aire 5 cops podem obtenir dos valors: cara, cara, creu, cara i creu.

7.2 VARIABLE ESTADISTICA

Una **variable estadística** es cada una de les **característiques o qualitats** que tenen els individus d'una població.

7.2.1 Tipus de variables

a) Variable qualitativa

Les **variables qualitatives** es refereixen a les característiques o qualitats que no poden ser mesurades amb un nombre. L'estat civil, color d'ulls són exemples de variables qualitatives.

b) Variable quantitativa

Una **variable quantitativa** és la que se expressa mitjançant un **nombre**, Amb elles es poden realitzar operacions aritmètiques. Podem diferenciar dos tipus:

Variable discreta

Una **variable discreta** és aquella que pren **valores aïllats** , es a dir **no** admet **valors intermitjos** entre dos valores específics. Per exemple el nombre de germans de 5 amics: 2, 1, 0, 3.

Variable continua

Una **variable continua** és aquella que pot prendre **valors compresos entre dos nom**

bres. Per exemple l'alçada de 5 amics: 1.73, 1.82, 1.77, 1.69, 1.75.

7.3 DISTRIBUCIÓ DE FRÈQÜENCIES

La **distribució de freqüències** o **taula de freqüències** és una ordenació en forma de taula de les dades estadístics, assignant a cada dada la seva freqüència corresponent.

7.3.1 Tipus de freqüències

a) Freqüència absoluta

La freqüència absoluta és el nombre de vegades que es repeteix un valor determinat, en un estudi estadístic. Es representa per n_i

La suma de totes les freqüències absolutes és igual al nombre de dades, que representarem per **N**

La suma de les freqüències absolutes es igual al número total de datos, que se representa por **N**. Per representar de forma resumida aquesta suma s'utilitza la lletra grega **Σ** (sigma majúscula) que es llegeix suma o sumatori.

$$\sum_{i=1}^{n} n_i = N$$

b) Freqüència relativa

La freqüència relativa és el quocient entre la freqüència absoluta (n_i) d'un determinat valor i el nombre total de dades (N). Es designa per $\boldsymbol{f_i}$

$$f_i = \frac{n_i}{N}$$

La suma de les freqüències relatives és igual a 1.

$$\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} \frac{n_i}{N} = 1$$

c) Freqüència absoluta acumulada

La freqüència absoluta acumulada és la suma de les freqüències absolutes de tots els valors inferiors o iguals al valor considerat. És representa per $\mathbf{N_i}$

d)Freqüència relativa acumulada

La freqüència relativa acumulada és la suma de les freqüències relatives de tots els valors inferiors o iguals al valor considerat. És representa per $\mathbf{F_i}$

EXEMPLE

Durant el mes de juliol, en una ciutat s'han registrat les següents temperatures màximes:

32, 31, 28, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31, 30, 30, 29, 29, 30, 30, 31, 30, 31, 34, 33, 33, 29, 29.

Organitzeu les dades en una taula de freqüències

En la primera columna de la taula indiquem tots els valor de la variable ordenats de menor a major, en la segona fem el recompte i en la resta les diferents tipus de freqüències:

x_i	Recompte	n_i	N_i	f_i	F_i
27	I	1	1	0.032	0.032
28	II	2	3	0.065	0.097
29	~~IIII~~ I	6	9	0.194	0.290
30	~~IIII~~ II	7	16	0.226	0.516
31	~~IIII~~ III	8	24	0.258	0.774
32	III	3	27	0.097	0.871
33	III	3	30	0.097	0.968
34	I	1	31	0.032	1
		31		1	

Aquest tipus de taules de freqüències s'utilitzen amb variables discretes

7.3.2 <u>Distribución de freqüències agrupades</u>

La distribució de freqüències agrupades o taula de dades agrupats s'utilitza si la variable pren un nombre de valors diferents molt gran o la variable és continua.

En aquests casos s'agrupen els valors en intervals que tinguin la mateixa amplitud i que denominarem classes. A cada classe se li assigna la seva freqüència corresponent.

- **Límits de la classe**

Cada classe està delimitada per el límit inferior de la classe i el límit superior de la classe.

- **Amplitud de la classe**

L'amplitud de la classe és la diferència entre el límit superior i inferior de la classe.

- **Marca de classe**

La marca de classe és el punt mig de cada interval, i és el valor que representa a tot l' interval per al càlcul de alguns paràmetres

- **Construcció de una taula de dades agrupades**

Donades les següents dades, construïu una taula de freqüències de dades agrupades

3, 15, 24, 28, 33, 35, 38, 42, 43, 38, 36, 34, 29, 25, 17, 7, 34, 36, 39, 44, 31, 26, 20, 11, 13, 22, 27, 47, 39, 37, 34, 32, 35, 28, 38, 41, 48, 15, 32, 13.

Passos a seguir:

1. Localitzar els valors menor i major. En aquest cas són 3 i 48
2. Calcular l'amplitud de les dades, restant al valor més gran el més petit
   48 – 3 = 45
3. Buscar un nombre sencer un poc superior a la diferencia i que sigui divisible pel nombre d'intervals que pretenem establir, Es convenient que el nombre d'intervals oscil·li entre 6 i 15.

   En aquest cas podem fer 10 intervals d'amplitud 5,

4. Es formen els intervals de manera que el límit inferior d'una classe pertany a l' interval, però el superior no formarà part d'aquesta classe si no de la següent.

	c_i	n_i	N_i	f_i	F_i
[0, 5)	2.5	1	1	0.025	0.025
[5, 10)	7.5	1	2	0.025	0.050
[10, 15)	12.5	3	5	0.075	0.125
[15, 20)	17.5	3	8	0.075	0.200
[20, 25)	22.5	3	11	0.075	0.2775
[25, 30)	27.5	6	17	0.150	0.425
[30, 35)	32.5	7	24	0.175	0.600
[35, 40)	37.5	10	34	0.250	0.850
[40, 45)	42.5	4	38	0.100	0.950
[45, 50)	47.5	2	40	0.050	1
		40		**1**	

7.4 PARÀMETRES ESTADÍSTIC

7.4.1 Definició de paràmetre estadístic

Un paràmetre estadístic és un nombre que s'obté a partir de les dades d'una distribució estadística.
Els paràmetres estadístics són molt útils per sintetitzar la informació donada per una taula o per una gràfica.

Hi ha dos tipus de paràmetres estadístics:

- De centralització.
- De dispersió.

a) Paràmetres de centralitzacions

Aquests paràmetres ens indiquen al voltant de quin valor (centre) es distribueixen les dades.

Els paràmetres de centralització són tres:

Mitjana aritmètica

La mitjana aritmètica és el valor promig de la distribució, és representa per $\bar{x}$ i es calcula:

$$\bar{x} = \sum_{i=1}^{n} x_i \cdot f_i = \frac{\sum_1^n x_i \cdot n_i}{N}$$

Mediana

La mediana és el valor que separa la meitat superior de la distribució de la inferior, es a dir que divideix la sèrie de dades en dos parts iguals

Si N és senar, només hi ha un terme central, la mediana és el valor que ocupa el lloc $\frac{N+1}{2}$

Si N és parell, hi ha dos termes centrals, en aquest cas la mediana serà la mitjana aritmètica dels valors que ocupen els llocs $\frac{N}{2}$ i $\frac{N+1}{2}$

Moda

La moda és el valor més freqüent de la distribució (els valor que es repeteix més cops)

EXEMPLE

Durant el mes de juliol, en una ciutat s'han registrat les següents temperatures màximes:

32, 31, 28, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31, 30, 30, 29, 29, 30, 30, 31, 30, 31, 34, 33, 33, 29, 29.

Trobeu la mitjana aritmètica, la mediana i la moda

x_i	n_i	$x_i \cdot n_i$
27	1	27
28	2	56
29	6	174
30	7	210
31	8	248
32	3	96
33	3	99
34	1	34
	$\sum_1^8 n_i = 31$	$\sum_1^8 x_i \cdot n_i = 994$

Mitjana aritmètica $\rightarrow \bar{x} = \frac{\sum_1^8 x_i \cdot n_i}{N} = \frac{994}{31}$

La temperatura mitjana del mes de juliol és de 30.45 graus

La mediana serà el valor que ocupa el lloc $\frac{31+1}{2} =$ 16 $\rightarrow$ La mediana és 30 graus

La moda serà 31 graus

b) Mesures de dispersió

Les mesures de dispersió ens informen de quant s'allunyen del centre els valors de la distribució.

Les mesures de dispersió són:

- Rang o recorregut
- Variància
- Desviació típica

Rang o recorregut

El rang o recorregut és la diferència entre el major i el menor valor d'una distribució estadística

Variància

La variància és la mitjana aritmètica del quadrat de les desviacions respecte de la mitjana aritmètica. Es representa per σ^2:

$$\sigma^2 = \frac{\sum_1^n (x_i - \bar{x})^2 \cdot n_i}{N} = \frac{\sum_1^n {x_i}^2 \cdot n_i}{N} - \bar{x}^2$$

Desviació típica

La desviació típica és l'arrel quadrada de la variància, es representa per σ

$$\sigma = \sqrt{\frac{\sum_1^n (x_i - \bar{x})^2 \cdot n_i}{N}} = \sqrt{\frac{\sum_1^n {x_i}^2 \cdot n_i}{N} - \bar{x}^2}$$

EXEMPLE

Durant el mes de juliol, en una ciutat s'han registrat les següents temperatures màximes:

32, 31, 28, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31, 30, 30, 29, 29, 30, 30, 31, 30, 31, 34, 33, 33, 29, 29.

Trobeu el rang, la variància i la desviació

x_i	n_i	$x_i^2 \cdot n_i$
27	1	729
28	2	1568
29	6	5046
30	7	6300
31	8	7688
32	3	3072
33	3	3267
34	1	1156
	$\sum_1^8 n_i = 31$	$\sum_1^8 x_i^2 \cdot n_i$ =28826

Rang o recorregut→ 34 – 27 = 7 graus

Variància →

$$\sigma^2 = \frac{\sum_1^n {x_i}^2 \cdot n_i}{N} - \bar{x}^2 = \frac{28826}{31} - 30.45^2 = 2.67$$

Desviació típica → $\sigma = \sqrt{2.67}$

PROBABILITAT

7.5 DEFINICIONS

La **probabilitat** de un succés és un nombre, comprés entre 0 i 1, que indica la possibilitat que té de verificar-se quan es realitza un experiment aleatori

Experiments deterministes són aquells que podem predir el resultat abans de que es realitzen.

Per exemple, deixem caure una pedra des de una finestra, sabem sense cap dubte, que la pedra baixarà. Si la llancem cap a dalt, també sabem amb tota seguretat que pujarà durant un interval determinat de temps, però desprès baixarà

Experiment aleatori són aquells que no es pot predir el resultat, ja que aquest depèn de l'atzar.

Per exemple si llancem una moneda no sabrem amb antelació si sortirà cara o creu. Si llencem un dau tampoc podem determinar el resultat que obtindrem

La **teoria de probabilitats** s'ocupa d'assignar un cert nombre a cadascun dels possibles resultats que se poden obtenir en un experiment aleatori, a fi de quantificar dits resultats i saber si un succés és més probable que un altre.

Introduïm una sèrie de definicions:

Un **succés** és cadascun dels resultats possibles d'una experiència aleatòria.

Al llançar una moneda obtenir cara
Al llançar un dau obtenir 4

Espai mostral és el conjunt de tots els possibles resultats d'una experiència aleatòria, es representa per E

Espai mostral de llançar una moneda → E = $\{cara, creu\}$
Espai mostral de llançar un dau → E = $\{1,2,3,4,5,6,7\}$

Succés aleatori és qualsevol subconjunt d'un espai mostral
Per exemple al llançar un dau un succés aleatori seria que surti parell, un altre, obtenir un múltiple de 3, un altre obtenir un 5

Succés elemental és cadascun dels elements que formen part del espai mostral.
Per exemple al llançar un dau un succés elemental és que surti un 5

Succés compost és qualsevol subconjunt de l'espai mostral
Per exemple al llançar un dau un succés compost és que surti un parell o múltiple de 3

Succés segur, E, està format per tots els possibles resultats (es a dir per tot l'espai mostral)
Per exemple al llançar un dau un succés segur és obtenir una puntuació menor que 7

Succés impossible, $\emptyset$ (conjunt buit), és aquell que no té ningun element.
Per exemple al llançar un dau un succés segur és obtenir una puntuació igual o major que 7 .

Successos compatibles: Dos successos ,A i B, són compatibles quan tenen algun succés elemental comú.
Per exemple si A es obtenir una puntuació par al llançar un dau i B es obtenir un múltiple de 3; A i B són compatible perquè el 6 és un succés elemental comú

Successos incompatibles: Dos successos ,A i B, són incompatibles quan no tenen cap succés elemental comú.
Per exemple si A es obtenir una puntuació par al llançar un dau i B es obtenir un múltiple de 5; A i B són incompatible perquè no tenen cap elemental comú.

Successos independents: Dos successos ,A i B, són independents quan la probabilitat de que succeeixi A no es veu afectada perquè hagi succeït B.
Per exemple al llançar dos daus els resultats són independents

Successos dependents: Dos successos ,A i B, són dependents quan la probabilitat de que succeeixi A es veu afectada perquè hagi succeït B.
Per exemple treure dues cartes d'una baralla, sense reposició, són successos dependents

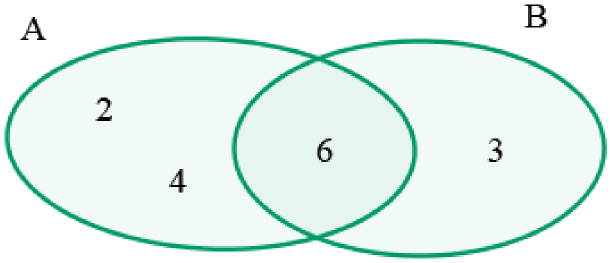
El **succés contrari** de A és un altre succés que es realitza quan A no es realitza, es designa per $\bar{A}$.
Dos successos contraris són traure senar i parell al llançar un dau

Espai de successos, S, és el conjunt de tots els successos aleatoris
Per exemple si llancem una moneda l'espai de successos està format per:

S = $\{\emptyset, \{c\}, \{x\}, \{c, x\}\}$
Observem que el primer element és el succés impossible i l'últim el succés segur

Si E té un nombre finit d'elements, n, el nombre de successos és 2^n.

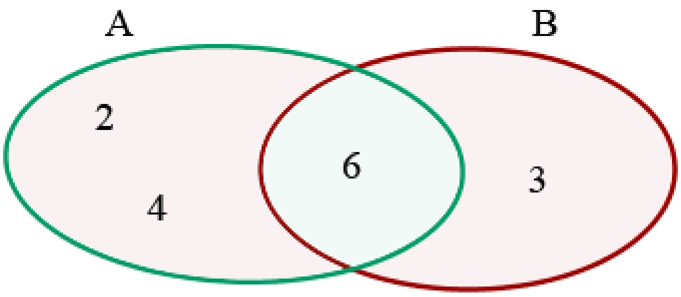
La **unió de successos, A∪B**, és el succés format per tots els elements de A i de B. Es a dir, el succés A∪B es verifica quan es verifica un dels dos A o B o tots dos. Es llegeix com A o B.
Per exemple l'experiment consisteix en llançar un dau, si A = "treure parell" i B = "treure múltiple de 3"
$A = \{2,4,6\}; B = \{3,6\} \rightarrow A \cup B = \{2,3,4,6\}$



La **intersecció de successos A∩B**, és els succés format per tots els elements que són alhora de A i B, es llegeix com A i B

Per exemple l'experiment consisteix en llançar un dau, si A = "treure parell" i B = "traure múltiple de 3"

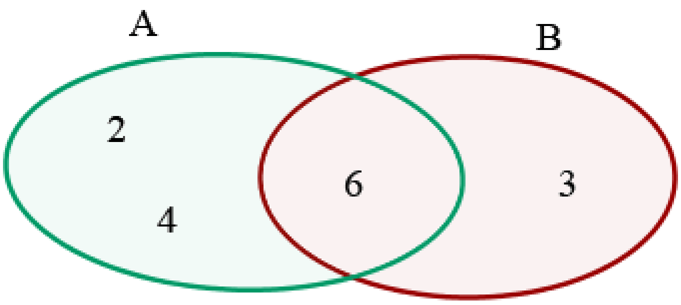
$A = \{2,4,6\}$; $B = \{3,6\}$ → A∩B= {6}



La **diferencia de successos A – B** , és el succés format per tots els elements de A que no són de B. Es llegeix A menys B

Per exemple l'experiment consisteix en llançar un dau, si A = "treure parell" i B = "traure múltiple de 3"

$A = \{2,4,6\}$; $B = \{3,6\}$ → A – B = {2,4}



7.6 AXIOMES DE LA PROBABILITAT

1. La probabilitat és positiva i menor o igual que 1 $0 \le p(A) \le 1$

2. La probabilitat del succés segur és 1 $p(E) = 1$

3. Si A i B són incompatibles, es a dir $A \cap B = \emptyset \rightarrow p(A \cup B) = P(A) + p(B)$

7.7 PROPIETATS DE LA PROBABILITAT

1. $p(A) + p(\bar{A}) = 1 \rightarrow p(\bar{A}) = 1 - p(A)$

2. $p(\emptyset) = 0$

3. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

4. Si $A \text{c} B \rightarrow p(A) \leq p(B)$

5. Si A_1, A_2 A_3............ A_n són incompatibles dos a dos:

$$p(A_1 \cup A_2 \cup A_{3 \ldots\ldots\ldots} \cup A_n) = p(A_1) + p(A_2) + p(A_3) + \cdots p(A_n)$$

6. Si E és finit i un succés $S = \{x_1, x_2 \ldots \ldots \ldots x_n\}$

$$p(S) = p(x_1) + p(x_2) + p(x_3) + \cdots p(x_n)$$

Per exemple la probabilitat de treure parell, al llançar un dau, és

P(par) = p(2) + p(4) + p(6)

7.8 CÀLCUL DE PROBABILITATS

1. **Regla de Laplace**: Si realitzem un experiment aleatori en el que hi ha n successos elementals tots iguals de probables, **equiprobables**, aleshores si A és un succés:

$$p(A) = \frac{nombre\ de\ casos\ favorables\ a\ A}{nombre\ de\ casos\ possibles}$$

Per exemple en una baralla de 40 cartes trobeu la probabilitat de treure un as i la probabilitat que la carta sigui de copes

$$p(as) = \frac{4}{40} = \frac{1}{10}; \quad p(copes) = \frac{10}{40} = \frac{1}{4}$$

2. Probabilitat de la unió de successos incompatibles

$$\text{Si } A \cap B = \emptyset \rightarrow p(A \cup B) = P(A) + p(B)$$

Exemple

Calculeu la probabilitat d'obtenir un 2 o un 5 al llançar un dau

A = obtenir un 2 ; B = obtenir un 5 → $A \cap B = \emptyset \rightarrow p(A \cup B) = P(A) + p(B)$

$$p(\mathbf{2} \cup \mathbf{5}) = P(\mathbf{2}) + p(\mathbf{5}) = \frac{\mathbf{1}}{\mathbf{6}} + \frac{\mathbf{1}}{\mathbf{6}} = \frac{\mathbf{2}}{\mathbf{6}} = \frac{\mathbf{1}}{\mathbf{3}}$$

3. Probabilitat de la unió de successos compatibles

$$Si\ A \cap B \neq \emptyset \rightarrow p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Exemple

Calculeu la probabilitat d'obtenir un múltiple de 2 o un 6 al llançar un dau

$A = \{2,4,6\}$; B = {6} → $A \cap B = \{\mathbf{6}\}$

$$p(múltiple\ de\ 2 \cup 6) = P(múltiple\ de\ 2) + p(6) - p(6) = \frac{3}{6} + \frac{1}{6} - \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

4.Siguin A i B dos successos d'un mateix espai mostral E. Anomenem **probabilitat condicionada** del succés B condicionat a A, i es representa per $p(B/_A)$ a la probabilitat del succés B un com ha passat A

$$p(B/_A) = \frac{p(A \cap B)}{p(A)}$$

Exemple

Calculeu la probabilitat d'obtenir un 6 al llançar un dau sabent que ha sortit un nombre parell

$$p\left(6/_{parell}\right) = \frac{p(6 \cap parell)}{p(parell)} = \frac{1/_6}{3/_6} = \frac{1}{3}$$

5.Probabilitat de la intersecció de successos independents

$$p(A \cap B) = p(A) \cdot p(B)$$

Exemple

Tenim una baralla de 40 cartes, es treu una carta i es torna aposar a la baralla. Quina és la probabilitat de extreure dos asos.

Donat que tornem la primera carta, la extracció segona no depèn de la primera per tant

$$p(A_1 \cap A_2) = p(A_1) \cdot p(A_2) = \frac{4}{40} \cdot \frac{4}{40} = \frac{1}{100}$$

6.Probabilitat de la intersecció de successos dependents

$$p(A \cap B) = p(A) \cdot p\left(\boldsymbol{B}/\boldsymbol{A}\right)$$

Exemple

Tenim una baralla de 40 cartes, es treuen dues cartes. Quina és la probabilitat de extreure dos asos.

Donat que no tornem la primera carta, la extracció segona depèn de la primera extracció per tant

$$p(A_1 \cap A_2) = p(A_1) \cdot p\left(A_2/A_1\right) = \frac{4}{40} \cdot \frac{3}{39} = \frac{1}{130}$$