**CURSUL 6**

**Serii statistice**

Obiectivele cursului

În studiul acestuia vizăm obținerea de cunoștințe privind:

* modalități de prezentare a datelor statistice;
* sistematizarea/gruparea datelor statistice
* utilizarea EXCEL în procesul de prezentare a datelor.

Enumerarea datelor statistice într-o anumită ordine se numeşte serie statistică***.*** Seria statistică se reprezintă ca două şiruri de date. Primul şir este criteriul de enumerare (valori , variante) al doilea şir conţine datele numerice.

Seriile statistice sunt de următoarele feluri

1. Serii de repartiţie sau de distribuţie (distribuţii empirice şi distribuţii teoretice)
2. Serii cronologice (ale dinamicii ) sau serii de timp.
3. Serii teritoriale sau de spaţiu.
4. Serii descriptive sau enumerative (lista candidaţilor admişi, facultatea, media de liceu, media examenelor, şcoala absolvită etc.)
5. Serii unidimensionale şi multidimensionale

La rândul lor seriile statistice pot fi atât numerice (cantitative) cât şi nenumerice (calitative).

Cu ajutorul graficilor se vizualizează informaţiile statistice facilitând perceperea pe ansamblu a datelor, aspecte privind: variaţia valorilor observate, repartiţia lor, legăturile existente între ele, a evoluţiei valorilor în timp ş.a.

Graficul trebuie să cuprindă:

1. Titlul – (precizându-se şi limitele fenomenului)
2. Legenda
3. Sistemul axelor rectangulare (ox, oy, oz)

Reprezentarile grafice ne permit sa prezentam datele introduse într-un format vizual folosind diferite tipuri de grafice (diagrame). Una dintre cele mai mari facilitati ale MS Excel este aceea de a construi reprezentări grafice. Graficele din MS Excel pot fi plane 2D sau in spațiu 3D şi fiecare tip de diagramă standard prezintă şi câteva subtipuri.

Tipuri de reprezentări grafice: grafice de volum, grafice de structură , grafice prin benzi, grafice prin coloane simple sau în aflux, cronogramă (historiogramă*)* pe scară uniformă sau logaritmică, diagrama polară (radială*),* histograma, poligonul frecvențelor, curba cumulativă a frecvențelor, graficul lui Lorentz, cartograma, cartodiagrama.

Cele mai utilizate sunt :

Grafice liniare afişează tendinţele din datele la intervale egale şi au urmatoarele subtipuri de diagrame: diagrame liniare plane 2D, diagrame liniare stratificate, diagramele liniare stratificate100%, diagramele liniare 3D

Diagramele coloana afișează modificările datelor pe o perioada de timp de moment sau ilustrează comparații între elemente. Momentele de timp sau categoriile sunt ordonate pe orizontală iar valorile pe verticală. Remarcăm : diagrame coloana grupata si grupata 3D, diagrame coloana stratificata si stratificata 3D, diagrama coloana stratificata 100% si 3D, diagramele con, cilindru sau piramidă (utilizează marcatori în forma de con, cilindru, piramidă pentru a conferi un aspect dramatic diagramelor de tip coloanăşi au aceleaşi subtipuri ca şi graficele coloană.

Prelucrarea statistica – este un proces prin care datele înregistrate sunt sistematizate si tratate statistic în vederea obţinerii sistemului de indicatori statistici. Obţinerea indicatorilor statistici este necesară pentru caracterizarea în expresii numerice generalizatoare, a ceea ce este esenţial pentru un grup de unităţi sau pentru întreaga colectivitate, studiate în condiţii de timp şi de spaţiu.

Prelucrarea primară înseamnă în acelaşi timp şi centralizarea, sistematizarea şi omogenizarea datelor observării în vederea aplicării modelelor de calcul şi analiză satistică

Sistematizarea datelor înregistrate presupune ordonarea acestora în funcţie de omogenitatea lor.

Deci la baza sistematizării se află principiul omogenităţii.

**Remarcă**

Principiul omogenităţii este un principiu de pierdere a informaţiei pentru a câştiga informaţie**.**

Sistematizarea se realizează prin : a) centralizarea şi b) gruparea statistica a datelor.

**CURSUL 7**

**Obiectivele cursului**

În studiul acestuia vizăm obținerea de cunoștințe privind:

* modalități de prezentare a datelor statistice;
* sistematizarea/gruparea datelor statistice

utilizarea EXCEL în procesul de prezentare a datelor.

1. Centralizarea – presupune totalizarea unităţilor statistice sau a valorilor unei caracteristici la nivelul grupelor tipice sau al colectivităţii observate. Totalizarea valorilor unei caracteristici se face prin însumarea directa sau prin folosirea unor coeficienţi de echivalenta (preturi, timpul de munca etc.) În urma centralizării obţinemindicatori statistici de nivel (ex, nr. populaţiei unei localităţi, valoarea producţiei)

* ***Centralizarea simplă***şi obţinerea indicatorilor (agregatelor) generali

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nr.crt.  al unităţii | Variabile însumabile direct | | | |
| **X1** | **X2** | **…** | **Xp** |
| 1 | *x11* | *x21* | *…* | *xp1* |
| 2 | *x12* | *x22* | *…* | *xp2* |
| … | … | … | … | … |
| i | *x1i* | *x2i* | *…* | *xpi* |
| … | … | … | … | … |
| n | *x1n* | *x2n* | *…* | *xpn* |
| Total |  |  | … |  |

Pe orizontală se găseşte fiecare unitate cu toate variantele înregistrate la toate caracteristicile (variabilele) înregistrate, iar pe verticală sunt distribuţiile de valori ale celor *p* variabile independente înregistrate.

Gruparea statistică– este o centralizare pe grupe a unităţilor unei colectivităţii. Se obţine în baza aplicării metodei grupărilor statistice presupunând principiul omogenităţii.

Omogenitatea statistică înseamnă ca în grupa (clasa) respectivă să existe variaţie minimă între variantele înregistrate – variaţie care poate fi explicată ca fiind influenţa factorilor aleatori.

În forma grupării statistice – rezultă şiruri de date ordonate după variaţie, în sens crescător sau descrescător, a caracteristicilor de grupare.

După conţinutul caracteristicii de grupare putem avea: a) grupări cronologice în cazul în care sistematizarea datelor se face după o variabilă de timp; b) grupări teritoriale când sistematizarea datelor se face după o variabilă de spaţiu; c)grupări atributivese folosesc pentru toate caracteristicile, în afara caracteristicilor de timp şi spaţiu.

Caracteristicile atributive pot fi caracteristici cantitative (numerice) şi calitative (nenumerice).

După numărul variabilelor de grupare pot fi: simple şi combinate***.***

Grupările simple**(**după o singură variabilă *)* pot fi:

1. gruparea pe variantedacă amplitudinea variaţiei este mică şi la nivelul unităţilor individuale s-au înregistrat un număr mic de valori distincte (variante);

|  |  |
| --- | --- |
| **Durata sejurului (zile )** | **Nr. turişti** |
| 5 | 10 |
| 6 | 15 |
| 7 | 20 |
| 8 | 30 |
| 9 | 35 |
| 10 | 30 |
| 11 | 15 |
| 12 | 10 |
| **TOTAL** | **165** |

1. gruparea pe intervale de variaţie egaledacă amplitudinea variaţiei este moderată. În acest caz, e necesar, să se stabilească numărul de grupe şi mărimea intervalului de variaţie.

|  |  |
| --- | --- |
| **Grupe de studenţi după nr.**  **de volume citite** | **Nr. studenţi** |
| 30- 50 | 50 |
| 50-70 | 80 |
| 70-90 | 60 |
| 90-110 | 120 |
| 110-peste | 130 |
| **Total** | **440** |

1. gruparea pe intervale de variaţie neegale*.* Prin gruparea pe intervale neegale se urmăreşte să se structureze colectivitatea pe tipuri calitative, dacă nu se cunosc valorile (pragurile) care separă tipurile calitative, se procedează mai întâi la gruparea pe intervale egale şi apoi se poate folosi criteriul mediei pentru formarea tipurilor calitative “mic”, “mediu”, “mare”. In acest caz se obţine o grupare tipologică.

|  |  |
| --- | --- |
| **Grupe de studenţi după nr. de ore aferente pregătirii pentru examene** | **Nr. studenţi** |
| Sub 30 | 4 |
| 30-40 | 5 |
| 40-60 | 6 |
| 60-85 | 10 |
| 85-120 | 12 |
| 120-150 | 7 |
| 150-200 | 4 |
| 200 peste | 2 |
| **Total** | **50** |

Gruparea combinată presupune sistematizarea datelor după două sau mai multe caracteristici de grupare (cel mult 4) care pot fi variabile numerice şi/sau calitative. Gruparea combinată impune stabilirea ordinii de grupare pe baza relaţiei de interdependenţă dintre variabile. Grupele formate după prima caracteristică se regrupează după cea de a doua ş.a.m.d.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Grupe după ore lucrate** | **Repartiţiile după veniturile săptămânale (lei)** | | | | **Total** |
| **Sub 1000** | **1000-2000** | **2000-3000** | **3000 şi peste** |
| sub 10 | 10 | 10 | 5 | - | **25** |
| 10 – 20 | - | 5 | 10 | - | **15** |
| 20 şi peste | - | - | 5 | 15 | **20** |
| **Total** | **10** | **15** | **20** | **15** | **60** |

Clasificarea **(**gruparea pentru variabile calitative*)*se efectuează după variabile nenumerice (calitative) şi presupune includerea în aceeaşi grupă (clasă) a tuturor unităţilor la care s-a înregistrat aceeaşi formă de manifestare a caracteristicii. De regulă, are un caracter oficial (ex. CAEN) şi în prealabil trebuie stabilit un nomenclator al claselor.

|  |  |
| --- | --- |
| **SEX** | **Nr studenţi** |
| FEMININ | 125 |
| MASCULIN | 200 |
| Total | 325 |

Gruparea pe intervale egale implică următoarele etape:

1. calculul amplitudinii absolutedevariaţiei (*A*) care exprimă împrăştierea maximă a valorilor serei.

*Axa = Xmax – Xmin;*

b) stabilirea în parametrii într-un anumit număr de grupe (r***)*** se poate stabili conform relaţiei lui Sturgers,

*r = 1+3,322log n*

unde: *n* - nr. caracteristicilor.

1. determinarea mărimii intervalului de grupare (h), se calculează ca raport între amplitudinea absolută a variaţiei şi numărul de grupe:

 sau 

Remarcă*.* Mărimea intervalului (*h*) se rotunjeşte la întreg în plus (ex.3,25≈4)

Prima grupă se porneşte de la *xmin* adăugându-se succesiv mărimea intervalului de grupare (*h*) rezultat din calculul anterior. Valoarea din stânga intervalului va fi considerată ca limită inferioară, iar valoarea din dreapta ca limită superioară. Valorile care formează limitele intervalelor se înregistrează fie numai ca limită superioară fie ca limită inferioară pentru a evita înregistrări duble a caracteristicilor.

Indiferent de scopul şi obiectul grupării, aceasta trebuie să îndeplinească cel puţin următoarele condiţii: a) completitudinea datelor; b) omogenitatea grupelor(claselor) şi subgrupelor (subclaselor); c) unicitatea includerii unităţilor într-o singură grupă (clasă); c) continuitatea variaţiei grupelor în cazul variabilelor numerice, ceea ce practic înseamnă să nu existe grupe cu frecvenţe nule care ar duce la întreruperea grupării.

Pentru analiza structurii colectivităţii pe grupe tipice se foloseşte gruparea pe intervale neegale.

Gruparea pe intervale neegale presupune regruparea intervalelor egale.

Un prim principiu al grupării pe intervale neegale este trecerea de la variaţia liniară (interval de

mărime constantă) la variaţia neuniformă a unor intervale de grupare din ce în ce mai mari. De exemplu, de la 8 grupe se poate trece la 5,3 grupe obţinând calificative .

Clasificarease efectuează după variabile nenumerice (calitative) şi are, de regulă, un caracter oficial (ex. CAEN) şi în prealabil trebuie stabilit un nomenclator al claselor.