

Treballs de recerca@

CÓM ES TRACTEN LES VARIABLES ESTADÍSTIQUES?

- TIPUS DE VARIABLES ESTADÍSTIQUES
- TIPUS DE VARIABLES QUE INTERVENEN EN UNA RECERCA EXPERIMENTAL
- TAULES DE FREQUÈNCIA
- REPRESENTACIONS GRÀFIQUES

Fonts bibliogràfiques:

- Estadística Bàsica. Universitat de Barcelona. 2000
- Anàlisi de dades. Universitat Ramón i Virgil. 2003
- Estadística: Modelos y métodos. Daniel Peña, Alianza Universidad Textos. 1994

TIPUS DE VARIABLES ESTADÍSTIQUES

Les variables estadístiques les classificarem segons el tipus de valors que poden prendre. Una primera divisió és:

- Variables estadístiques qualitatives (o nominals).
- Variables estadístiques quantitatives.

Variables estadístiques qualitatives són aquelles que no prenen valors numèrics.

Exemples: el color dels ulls, el sexe, el tipus de distracció preferit.

Variables estadístiques quantitatives són aquelles que prenen valors numèrics.

Dintre d'aquestes últimes encara podem fer una subdivisió:

- Variables estadístiques *quantitatives discretes*.
- Variables estadístiques *quantitatives contínues*.

Variable estadística és quantitativa discreta quan entre dos valors qualsevols que pot prendre la variable, només hi ha un nombre finit de possibles valors de la variable.

Exemples: el nombre de germans, el nombre de cotxes que passen en un dia per cert punt, el nombre de trucades que es reben cada hora en un ciutat.

Variable estadística és quantitativa contínua quan entre dos valors qualsevols que pot prendre la variable, hi pot haver un nombre infinit de possibles valors de la variable. Aquesta definició és equivalent a dir que si agafem dos valors que pot prendre la variable, tant propers com vulguem, sempre és possible trobar un altre valor de la variable que estigui entre els dos valors anteriors.

Exemples: l'alçada de les persones, el pes de les taronges, el temps que es tarda en fer un examen. S'ha d'observar, per exemple en el primer cas, que si agafem les altures de dues persones sempre és possible trobar-ne una altra que tingui una alçada entre les dues anteriors.



TIPUS DE VARIABLES QUE INTERVENEN EN UNA RECERCA EXPERIMENTAL

Moltes recerques científiques intenten reproduir i simular un determinat fenomen per tal de respondre una pregunta concreta. Aquest procés de simulació en condicions controlades rep el nom d'**experiment**.

Suposem que volem investigar quina influència té la llum sobre la germinació d'una determinada espècie de planta. Podríem formular la següent **pregunta**: *Com influeix la llum en la germinació de les llavors d'aquesta planta?*

El **fenomen** seria la *germinació*. Un grup d'alumnes formularia la **hipòtesi** que *potser les llavors d'aquesta planta germinaran abans en condicions de foscor (una hipòtesi igualment plausible podria ser justament la contrària: potser les llavors d'aquesta planta germinaran abans si estan exposades a la llum)*.

En tot experiment diferenciem com a mínim una variable independent i una variable dependent. La **variable independent** és aquella sobre la qual l'investigador intervé, i per tant és canviada premeditadament, mentre que la **variable dependent** és la que s'observa (i sovint es mesura).

En l'exemple anterior (Com influeix la llum en la germinació de les llavors d'aquesta planta?), la **variable independent** seria la *llum* (l'investigador canvia les condicions de llum) i la **dependent** és justament el fenomen de *germinació* (observat a partir de la presència i mida dels brots o del trencament de les cobertes de la llavor).

També en tot experiment cal **fixar** totes les altres variables que intervenen en el procés per tal que no interfereixen o emmascarin l'efecte de la variable independent.

En l'exemple anterior (Com influeix la llum en la germinació de les llavors d'aquesta planta?), caldrà *fixar* variables com ara la quantitat d'aigua, la temperatura o el substrat on estiguin les llavors, d'aquesta manera assegurem que investiguem tan sols la influència de la llum sobre la germinació.

Per últim cal tenir en compte que qualsevol procés experimental està sotmès a la pròpia variació del fenomen que s'investiga. Tan sols quan els resultats presentin regularitat en repetides ocasions podrem acceptar la fiabilitat de les conclusions. Per això és important fer **rèpliques** en els experiments científics.

En l'exemple de les llavors, es tractaria d'experimentar amb *moltes llavors*, evidentment de la mateixa espècie (però millor d'exemplars diferents).



TAULES DE FREQUÈNCIA

Si observem una variable estadística sobre un conjunt d'individus, obtindrem una sèrie de dades (que poden estar repetides o no). Aquestes dades, si són quantitatives, les podem ordenar. Si són qualitatives l'ordenació és arbitrària. Els valors de les dades ordenades els notarem x_1, x_2, \dots, x_k , on x_1 és el valor de la dada més petita, x_2 de el valor de la segona i així successivament.

Definició: la freqüència absoluta d'un valor és el nombre de vegades que apareix aquest valor a la sèrie.

Definició: la freqüència relativa és el quocient entre la freqüència absoluta del valor i el nombre total de dades de la sèrie.

Definició: la freqüència absoluta acumulada d'un valor és la suma de totes les freqüències absolutes dels valors de la sèrie des del principi fins el propi valor.

Definició: la freqüència relativa acumulada d'un valor és la suma de totes les freqüències relatives dels valors de la sèrie des del principi fins el propi valor.

Exemple: s'ha llançat un dau 20 vegades i s'ha obtingut el següent resultat: 2, 3, 6, 6, 2, 4, 4, 4, 2, 5, 2, 4, 2, 4, 2, 5, 1, 6, 2 i 2. Fem la taula de freqüències:

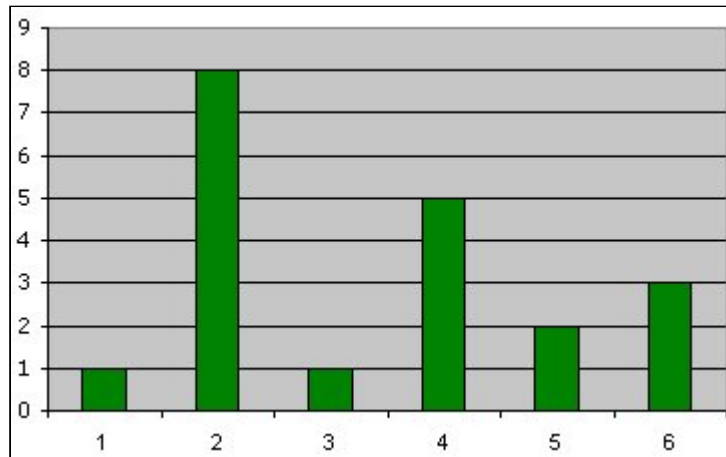
| Valor (x_i) | Freq. abs. (n_i) | Freq. rel. (f_i) | Freq. abs. ac. (N_i) | Freq. rel. ac. (F_i) |
|-----------------|----------------------|----------------------|--------------------------|--------------------------|
| 1 | 1 | 0.05 | 1 | 0.05 |
| 2 | 8 | 0.4 | 9 | 0.45 |
| 3 | 1 | 0.05 | 10 | 0.5 |
| 4 | 5 | 0.25 | 15 | 0.75 |
| 5 | 2 | 0.1 | 17 | 0.85 |
| 6 | 3 | 0.15 | 20 | 1 |



REPRESENTACIONS GRÀFIQUES

Diagrama de barres: l'usarem quan els valors de la variable estadística estiguin donats de manera individual. Normalment representarem les freqüències absolutes, les freqüències relatives o el tant per cent.

Exemple: amb les freqüències absolutes amb les dades de l'exemple anterior:



Histograma: l'usarem quan els valors de la variable estadística estiguin donats en forma d'interval. En aquest cas suposarem que els diferents intervals tenen la mateixa amplada. L'única diferència amb el diagrama de barres és que a l'eix horitzontal hem de posar els valors dels extrems dels intervals de manera ordenada i el gràfic es construeix aixecant sobre cada interval un rectangle vertical d'alçada igual a la freqüència que vulguem representar.

Exemple: histograma per representar les freqüències relatives les dades agrupades en intervals corresponents als pesos (en kg) de 16 alumnes d'una classe de secundària.

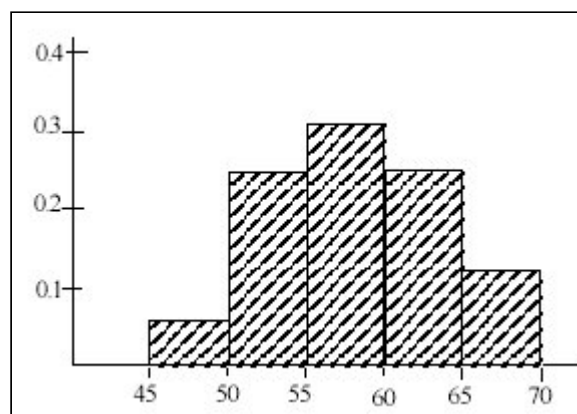


Diagrama de Pareto: aquest diagrama serveix per a representar dades qualitatives i es construeix ordenant les categories o classes per la seva freqüència relativa. S'utilitza molt en control de qualitat i permet detectar quin són els problemes més significatius en un procés (p.e. defectes en peces industrials, errors en l'enviament de productes...). Cada categoria es representa per un rectangle d'alçada igual a la seva freqüència relativa. En general el 80% dels problemes els produeixen el 20% de les possibles causes.

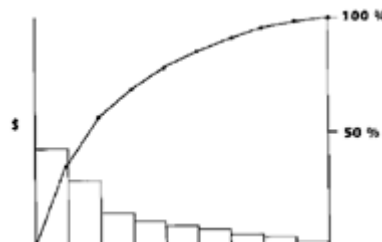


Diagrama de sectors: també conegut com diagrama de pastís per traducció literal de l'anglès "piechart". Donat un cercle de radi arbitrari, a cada valor es fa correspondre una porció d'àrea marcada per un determinat angle, que és proporcional a la seva freqüència relativa. L'àrea total són 360°.