


**Conferència pronunciada pel professor Pere Grima
de Facultat de Matemàtiques i Estadística de la UPC,
en la Jornada de l'ABEAM de l'11 de novembre de 2006.**

**“Estadístiques”, Enquestes i Experiments.
Algunes idees**

Pere Grima
pere.grima@upc.edu

Bon dia. La meva intenció és donar una visió general d'alguns aspectes de l'estadística que considero interessants, naturalment en un to de divulgació i sense aprofundir massa en res. Espero que, com a mínim, passem una hora entretinguda i, si a més, podeu agafar idees que us siguin útils, doncs molt millor.


Contingut

Primera capa: Estadística descriptiva
Mesurar la variabilitat
Percentatges
Gràfics

Segona capa: Mostreig
Conceptes clau. Males interpretacions
Grandària de mostra
Més enllà dels càlculs...

Tercera capa: Experiments
Experimentació i estadística?
Alguns exemples

2. "Estadístiques", enquestes i experiments. Algunes idees



He ordenat aquesta xerrada en 3 parts que he anomenat “capes” perquè crec que reproduïx un camí d'apropament a l'estadística. La primera capa, el primer contacte que es té, o que crec que s'hauria de tenir, és amb l'estadística descriptiva: mitjanes, gràfics... Això que popularment s'anomenen “estadístiques”. La segona capa és segurament el tema més tòpic, el tema amb que més s'identifica l'estadística a nivell general, que són les enquestes. I la tercera, una mica més profunda, és el disseny d'experiments de la qual en parlarem al final.


Contingut

Primera capa: Estadística descriptiva
Mesurar la variabilitat
Percentatges
Gràfics

Segona capa: Mostreig
Conceptes clau. Males interpretacions
Grandària de mostra
Més enllà dels càlculs...

Tercera capa: Experiments
Experimentació i estadística?
Alguns exemples

3. "Estadístiques", enquestes i experiments. Algunes idees



Comencem amb la primera capa: l'estadística descriptiva, centrant-nos en la importància de tenir en compte la variabilitat, usar correctament els percentatges i els gràfics.

Començant: Síntesi numèrica de dades

94,95	94,95	91,52
99,91	99,91	95,52
63,90	63,90	126,74
104,21	104,21	79,08
68,72	68,72	130,79
71,26	71,26	109,62
85,39	85,39	65,19
77,30	77,30	126,02
121,43	121,43	137,61
91,52	91,52	63,90
95,52	95,52	94,95
126,74	126,74	99,91
79,08	79,08	63,90
130,79	130,79	104,21
109,62	109,62	68,72
65,19	65,19	71,26
126,02	126,02	85,39
137,61	137,61	77,30
63,90	63,90	121,43
85,39	85,39	91,52
77,30	77,30	95,52
121,43	121,43	126,74
91,52	91,52	79,08
95,52	95,52	130,79
126,74	126,74	109,62
79,08	79,08	65,19
130,79	130,79	126,02
109,62	109,62	137,61
65,19	65,19	63,90
126,02	126,02	
137,61	137,61	
63,90	63,90	

SÍNTESI NUMÈRICA

Mesures de tendència central
Mitjana, Mediana i Moda

Mesures de dispersió
 Rang, Variància, **Desviació tipus ?**

Percentatges

4. "Estadístiques", enquestes i experiments. Algunes idees

Si tenim un gran volum de dades podem resumir la informació que contenen en només uns pocs valors i això, naturalment, pot ser molt útil. Aquest és un primer objectiu de l'estadística.

Quins són aquests valors? Bàsicament es divideixen en dos grans grups: en les mesures de tendència central (allò de "mitjana, mediana i moda...") i en les mesures de la dispersió (variància, desviació tipus...). També hi ha mesures anomenades de posició (quartils, percentils...) però no parlaré d'aquestes sinó dels percentatges, que també apareixen en aquests contextos i crec que mereixen una mica d'atenció.

D'aquestes mesures, les més importants, o les més útils si n'haguéssim d'escollir només dues, serien la mitjana i la desviació tipus. Tothom entén perfectament el que és la mitjana (els estudiants la practiquen quan calculen quines notes han de treure per aprovar). Però moltes vegades s'ignora, o no s'entén massa bé, el que significa la desviació tipus i és que, segurament, és un concepte molt més difícil d'interpretar que el de mitjana.

Però no convé ignorar-la. Si descrivim dades només amb mitjanes diríem coses com que "si un senyor es menja un pollastre sencer i un altre no en menja cap, en promig s'han menjat mig pollastre cada un", o que "si anem a la cuina de casa i posem els peus a la nevera i el cap al forn, se'ns queda el cos a la temperatura mitjana ideal".

Quantificar la variabilitat no és fàcil

La mitjana de les alçades es 170 cm
 Quina és la desviació tipus?

6. "Estadístiques", enquestes i experiments. Algunes idees

El problema crec que està en que quantificar la variabilitat, entendre el significat de la desviació tipus, no és intuïtiu i per això tampoc és gens fàcil.

Per exemple, si a un grup d'estudiants els hi pregunto quina és l'alçada mitjana de les persones adultes, tothom opinarà i en general estaran d'acord en que deu estar al voltant de 1,70m (curiosament, si li preguntes a una persona alta diu valors més aviat grans, i si li preguntes a una baixa els diu més aviat petits, com si la gent pensés que el valor d'un mateix

afecta de forma significativa a la mitjana general). Bé, és evident que no tothom mesura el mateix, i hem vist que ens convé ser capaços de mesurar i descriure la variabilitat, però si pregunto a estudiants que coneixen bé el que és la desviació tipus, i que l'han calculat moltes vegades, quina és la desviació tipus de les alçades, ja no tothom té les idees tan clares.

Si la variabilitat respon al patró de la distribució Normal, cas molt freqüent i que podem considerar que s'adapta bé a la nostra situació, podem fer servir les propietats que ens diuen que:

- a l'interval mitjana \pm 1 desviació tipus tenim el 68% de les observacions.
- a l'interval mitjana \pm 2 desviacions tipus tenim el 95 % (en realitat, és en l'interval mitjana \pm 1,96 desviacions tipus però habitualment s'arrodoneix a 2. Si posem 2 desviacions l'interval resultant és del 95,5%. Després veurem que això surt quan parlem de resultats d'enquestes).
- i a l'interval mitjana \pm 3 desviacions tipus tenim el 99,7% de les observacions.


És a dir, més enllà de la mitjana més tres desviacions tipus tenim en torn a l'1 o el 2 per mil d'observacions. I això és independent de quins siguin els valors de la mitjana i de la desviació tipus.

On posem el límit superior de les alçades? Segurament a 1,90m, o 1,95m. Per tant, la desviació tipus deu estar entre 7 i 8cm (si fos 7cm les 3 desviacions tipus estarien a 1,91m, i si fos 8cm a 1,94m).

"Els marcians existeixen!"

Els marcians existeixen!
Quina alçada tenen?

Mitjana = 150 cm
Desv. tipus = 20 cm



¿Quina alçada tenen els seus jugadors professionals de basquet?

6. "Estadístiques", enquestes i experiments. Algunes idees

Els marcians existeixen! Quina alçada tenen? Si la mitjana és 1,50m ¿vol dir que són més baixets que nosaltres? No necessàriament, hi podrien haver marcians de 10cm i marcians de 5m.

Suposem que la desviació tipus és de 0,2m i la distribució de les dades és similar a la que tenim els humans. Si fessin un partit de bàsquet terrícoles contra marcians (hauríem de suposar algunes coses més....), qui guanyaria? Tenint en compte que els que juguen a


bàsquet són els més alts, i que a la Terra mesuren al voltant de 2,10m o 2,15m, això vol dir que –suposant que la desviació tipus són 7cm- estan a unes 6 desviacions tipus de la mitjana. Si els marcians també posen als més alts, tindran una alçada de $1,5+6*0,2 = 2,7m$, i segur que ens guanyen.

Entendre la variabilitat és important...

Perill!: Els percentatges

Aparentment son trivials però les males interpretacions son freqüents:

- **Percentatge... de què?:**
Increment sobre el preu inicial o sobre el preu final. Regalen el 33% de producte (o el 25%?)
- **Percentatges i punts percentuals**
Els beneficis han passat de 2 al 4%. Han augmentat 2 punts ("un 2%") o han augmentat un 100%.



7. "Estadístiques", enquestes i experiments. Algunes idees

I ara passem a un tema molt més fàcil: els percentatges. Segurament molts pensaran que és trivial. De fet, això no s'acostuma a explicar en el context de l'estadística, tot i que també és una forma de resumir informació numèrica.

Crec que tot i ser trivials, moltes vegades ens equivoquem en fer-los servir, i cal estar atent a com s'utilitzen.

Per exemple: Si el contingut anterior d'un producte era 75ml i ara és de 100ml, quant ens regalen? La resposta no és evident. Si ens referim al contingut inicial regalen el 33%, però del contingut total que estan venent ara, només regalen el 25%. Davant d'un percentatge sempre cal preguntar-se: percentatge, de què? I descobrirem que moltes vegades no ho té clar ni el que ho ha calculat.

També hi ha confusions entre percentatges i punts percentuals (els beneficis han passat del 2% al 4%, han augmentat 2 punts percentuals o un "2%?"), percentatges sobre nivells o sobre canvis de nivell, operacions amb percentatges...

En els pòsters sobre els mals usos de l'estadística també hi ha algun comentari sobre això.

Percentatges. Exemple

El 65% dels menors de 10 a 17 anys admeten que accedeixen a programes per a majors d'edat

El Periódico, 05 gener 2006. Pàgina 27

Fins i tot amb premisses tan poc edificants, els adults poden fer el que vulguin. «El problema és que el 50% dels nens i el 15% de les nenes d'entre 10 i 17 anys reconeixen que fan servir habitualment videojocs»

8. "Estadístiques", enquestes i experiments. Algunes idees

Només un exemple que vaig veure fa uns mesos al diari: “El 65% dels menors...” però d'on surt aquest 65%?: del 50% dels nens i el 15% de les nenes!

I si sortís del 100% dels nens i el 100% de les nenes, tindriem aleshores el 200% dels menors...?

Començant: Síntesi numèrica de dades

94,95	94,95	91,04
89,91	95,91	95,52
63,65	63,65	126,74
104,21	104,21	79,08
68,72	68,72	130,79
71,26	71,26	109,62
85,39	85,39	65,19
77,30	77,30	126,02
121,43	121,43	137,61
91,52	91,52	63,90
95,52	95,52	94,95
126,74	126,74	99,91
79,08	79,08	63,65
130,79	130,79	104,21
109,62	109,62	68,72
65,19	65,19	71,26
126,02	126,02	85,39
137,61	137,61	77,30
63,90	63,90	121,43
85,39	85,39	91,52
77,30	77,30	95,52
121,43	126,74	
91,52	79,08	
95,52	130,79	
126,74	109,62	
79,08	65,19	
130,79	126,02	
109,62	137,61	
65,19	63,90	
126,02	137,61	
137,61	63,90	

GRÀFIQUES
Diagrames bivariants, histogrames...

Scatterplot of Weight vs Height

9. "Estadístiques", enquestes i experiments. Algunes idees

Passem ara a parlar una mica de les gràfiques que també són una excel·lent manera –crec que la millor– de resumir, i també de captar, la informació que contenen les dades.

Anàlisi exploratòria de dades

Resultat de la "Course Navette", pulsacions, alçada, pes, sexe, pràctica habitual d'algun esport,...

1	4	8
3	5	44
6	5	888
24	6	000012222222224444
40	6	666688888888888888
(17)	7	000002222222444444
35	7	6666688888
25	8	0002224444
15	8	67888
10	9	0000224
3	9	66
1	10	0

Scatterplot of Weight vs Height

10. "Estadístiques", enquestes i experiments. Algunes idees

Per exemple, crec que els gràfics es poden utilitzar per relacionar característiques referents amb l'estat de forma o el rendiment físic, i característiques anatòmiques com el pes, l'alçada, o si es practica o no algun esport de forma regular, si es fuma o no, etc... que poden recollir els alumnes sobre ells mateixos.

Dintre de l'àmbit acadèmic, però a un nivell més “professional”, un estudiant de la diplomatura d'estadística, una mica singular perquè ja era metge quan va començar la diplomatura, va fer el seu projecte final de carrera sobre l'establiment d'uns criteris per identificar possibles casos d'asma infantil no diagnosticats, a partir dels resultats de la “course navette” relacionats amb característiques físiques o anatòmiques de l'individu. Tot i que pot donar falsos positius (sembla que no pugui córrer, o que per les seves condicions hauria de córrer més) no sent realment un problema respiratori, sinó que podria ser

que simplement aquell dia no tingués ganes de córrer. Tot i així, d'una manera molt barata (de fet, no es fa res que no es faria normalment a la classe d'educació física) es poden identificar casos d'asma que d'altra manera podrien empitjorar i després fer més difícil el tractament i la recuperació.

Tornem a l'àmbit estrictament "escolar". Aquests gràfics els he fet a partir d'unes dades que inclou el paquet de software estadístic MINITAB i que explica que van ser obtingudes en una classe on hi havia 92 estudiants. Primer tots es van prendre les pulsacions i després van llençar una moneda, de manera que als que els va sortir cara van sortir a córrer. Quan van tornar es van prendre de nou, tots ells, les pulsacions. No és segur que les dades estiguin bé. De fet, quasi segur que no ho estan. Resulta que només els va tocar córrer a 35 persones (de 92) i és fàcil calcular que la probabilitat que això passi és menor del 5%. Tampoc és veritat que estiguessin 1 minut prenent-se les pulsacions -gairebé totes les dades són parells-, segurament van estar 30 segons (o 15) i les van multiplicar per 2 o per 4, i això és rellevant a efectes de l'error que es comet. Un error al comptar una pulsació (molt fàcil de cometre) es multiplica per 4 si ho fem per aquest procediment de comptar només 15 segons i després multiplicar.

També s'inclouen dades sobre hàbits i característiques físiques d'aquests estudiants. Es pot veure que les característiques d'alçada i pes són diferents pels homes i les dones.

Crec que aquí, en una classe d'estadística, no ens atreviríem a fer-los sortir a córrer per després mesurar-los l'increment de pulsacions. Però sembla que als Estats Units això sí que es pot fer.

Histograma "viu"



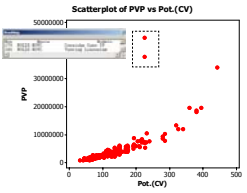
11. "Estadístiques", enquestes i experiments. Algunes idees

Fins i tot he trobat que fan coses com aquesta: "Living histogram". Un histograma format per les mateixes persones. Alts i baixos.

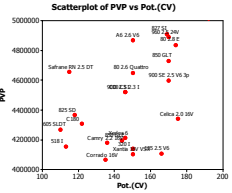
Cotxes, motos, ...

Relació entre preu i prestacions.
Quins són especialment cars o barats?

Scatterplot of PVP vs Pot.(CV)



Scatterplot of PVP vs Pot.(CV)



12. "Estadístiques", enquestes i experiments. Algunes idees

També donen molt joc, i crec que agrada a molts joves, les dades sobre característiques tècniques de cotxes (o també de motos) que són fàcils de trobar a internet o en revistes especialitzades, i que permeten estudiar com varia el consum dels automòbils en funció del pes, o l'acceleració en funció de la potencia o de la cilindrada, etc. També quins cotxes són singularment barats o cars en funció de les seves característiques...

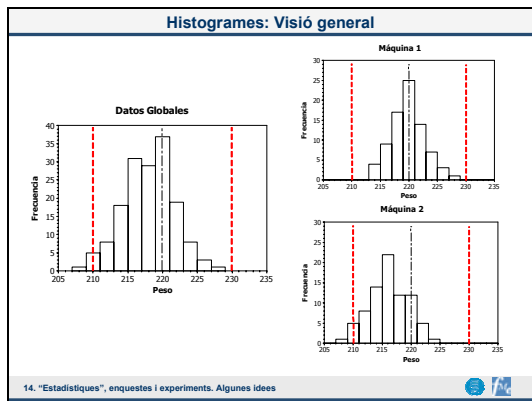
Aquests gràfics s'han fet amb unes dades ja una mica antigues, però és veu clarament que els Rolls Roys tenen un comportament molt singular, i això crec que no ha canviat.

Gràfiques per analitzar dades

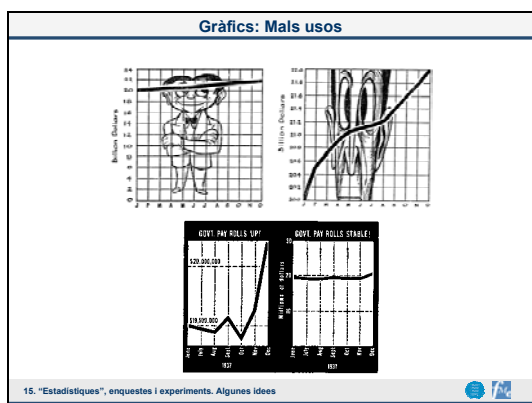
Dia	Ope.	Màquina 1				Màquina 2			
1	A	220.3	215.5	219.1	219.2	220.3	208.0	214.4	219.2
2	B	215.8	222.0	216.9	213.6	216.9	213.4	217.7	217.7
3	B	220.4	218.7	216.6	219.6	222.9	219.7	209.4	221.6
4	B	221.5	227.0	219.5	222.5	223.1	215.3	220.4	215.6
5	A	215.7	225.3	223.0	218.0	216.0	210.9	221.4	210.9
6	A	222.7	215.1	219.6	217.3	212.1	213.0	218.0	216.5
7	A	218.0	218.3	217.9	213.0	215.9	218.0	213.5	219.2
8	B	219.4	218.3	216.7	224.1	216.2	218.4	218.6	214.9
9	B	219.8	222.6	219.1	217.7	216.2	212.2	216.9	214.9
10	A	220.2	219.5	222.4	219.9	222.9	214.3	219.1	216.7
11	B	218.0	223.9	219.6	221.9	214.9	212.6	219.4	213.3
12	B	219.3	219.5	218.8	219.9	219.0	216.7	216.4	213.5
13	B	220.0	214.1	224.3	217.4	218.0	219.5	219.5	222.3
14	A	223.9	220.8	219.5	219.6	211.8	218.2	218.3	217.4
15	A	218.1	218.8	218.4	217.9	214.6	215.7	218.0	216.4
16	B	216.9	221.6	220.6	222.6	215.6	220.4	217.3	216.2
17	B	217.9	225.7	222.2	215.1	212.5	214.6	209.7	211.3
18	A	224.2	216.2	219.9	220.4	215.8	219.9	216.5	211.9
19	A	214.1	219.7	222.4	224.5	213.7	209.7	216.9	213.1
20	A	221.1	225.0	222.7	222.2	212.5	217.5	217.4	215.7

13. "Estadístiques", enquestes i experiments. Algunes idees

Els gràfics també serveixen per analitzar dades i treure conclusions clares i fàcils d'explicar, sense necessitat d'utilitzar tècniques sofisticades. Per exemple, aquestes dades es van obtenir en un forn de pa en el qual tenien dues màquines i hi treballaven dos operaris. Les barres de pa havien de pesar $220g \pm 10g$, i el propietari estava preocupat perquè sospitava que algunes barres estaven fora de "toleràncies".

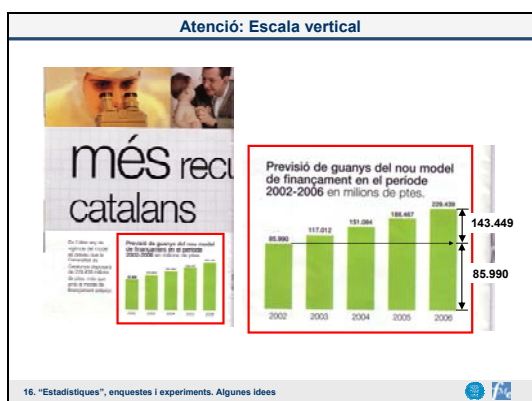


Si fem un histograma de totes les dades veiem que efectivament hi ha una tendència a que algunes unitats surtin fora dels límits establerts. I si estratifiquem per màquina (les dades es van recollir de manera que aquesta estratificació es pogués fer) s'observa que el problema està en la màquina 2, que estava descentrada, fabricant al voltant de 215g i no de 220g.



També es parla molt sobre la utilització dels gràfics de la forma que convingui per veure només allò que interressi. Per exemple, col·locant convenientment les escales.

El gràfic de l'esquerra dóna la sensació de que les despeses han pujat però d'una manera molt suau, mentre que el de la dreta, amb les mateixes dades, però fent un "zoom" sobre la part de l'escala on està la gràfica, l'increment es veu molt més exagerat.

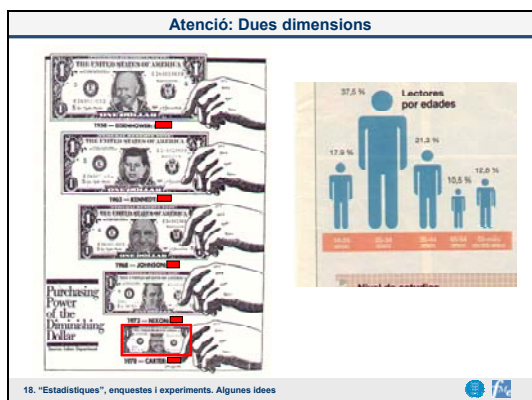


Els exemples anteriors són de llibre, però també se'n troben als diaris o als fulletons publicitaris. Per exemple, aquest és d'un fulletó que va editar la Generalitat on explica la previsió de finançament per al període 2002-2006. No sé com ha quedat això ara que ja estem acabant aquest període, però el que està clar és que aquest gràfic està mal fet, i l'increment respecte als primers anys és molt més gran del que el gràfic suggereix.



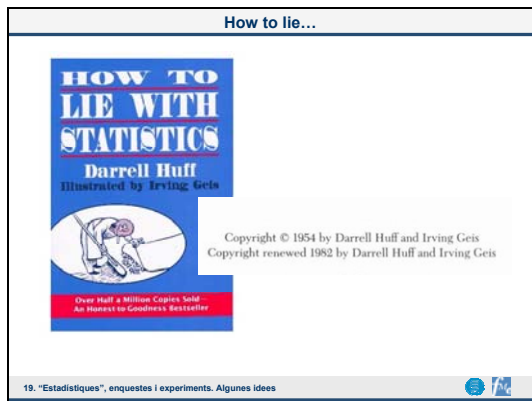
També cal anar amb compte amb l'eix horitzontal. Aquesta notícia sobre que "l'Ajuntament només multa..." ve acompanyada d'un gràfic on es veu que, efectivament, cada any es posen menys multes. Però si ens hi fixem, l'any 1999 s'en van posar 6.278, l'any 2000 se'n van posar força menys però també diu que va ser un any en que la Guardia Urbana va fer vaga, i l'any 2001, que és el que sembla que justificaria el titular, només comprèn fins el mes d'abril! Si en aquestes dates fem una estimació per tot l'any podríem multiplicar per 3 i sortiria de l'ordre de 7.500, més que l'any 1999 i per tant sembla que no cal amoïnar-se pel ritme en que es van posant les multes.

Per tant, l'escala de l'eix horitzontal, evidentment, cal que sigui homogènia.



Un altre tema són els gràfics en dues dimensions, que si no estan ben construïts donen una impressió que no es correspon amb les dades que volen representar. Aquí tenim un exemple de llibre: Quant més valia un dòlar en l'època d'Eisenhower respecte l'època de Carter? Segurament esteu pensant que 4 o 5 vegades, però no, només val una mica més del doble. De fet, es pot comprovar que les longituds són proporcionals als valors, però els bitllets tenen dos dimensions, i la sensació de quantitat que donen és proporcional a la seva superfície, i no a la longitud. Podem veure que la superfície del d'Eisenhower és 5 vegades més gran que l'altra.

I en el gràfic de la dreta, tret d'un diari que fa una "estadística" sobre les edats dels seus lectors, tenim un exemple similar. Entre 25 i 34 anys tenen un 37,5% de lectors, i de més de 55 tenen un 12,8%. Tot i que aquests últims són 1/3 dels primers, en el gràfic sembla que no pinten quasi bé res. De fet, es pot posar "un home de 12,8%" a cada pota del "gegant". Han fet els gràfics proporcionals a les alçades (el "gegant" és tres vegades més alt que l'home que comprèn el grup de lectors de més de 55 anys), però la sensació que donen és proporcional a les àrees.



19. "Estadístiques", enquestes i experiments. Algunes idees

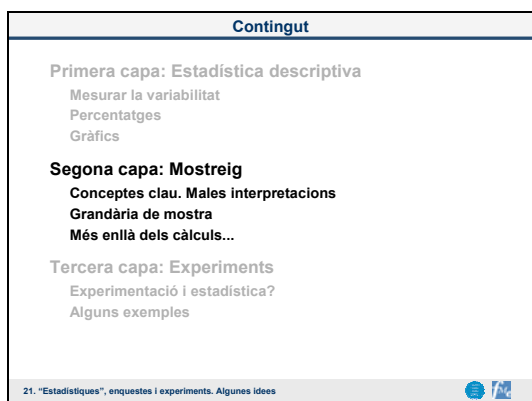
Si us interessa aquest tema de com enganyar o, potser millor dit, com evitar ser enganyat amb l'estadística, hi ha llibres molt interessants. El més famós és aquest titulat "How to Lie with Statistics", que no és un llibre nou, sinó que va sortir l'any 1954.



20. "Estadístiques", enquestes i experiments. Algunes idees

Abans de començar aquesta xerrada, quan instal·lavem l'ordinador, m'he connectat a internet i he anat a la pàgina d'Amazon, segurament la llibreria més gran del món, i hi he buscat els llibres que contenen la paraula "statistics" al seu títol, i després els he ordenat per vendes.

Surten 26.500 llibres aproximadament, i aquests que veiem són els més venuts. Sembla que hi ha demanda de llibres que expliquin estadística sense masses complicacions. El "How to lie..." que té més de 50 anys, l'he trobat alguna vegada en el número 1. Avui està en el 6, sobre més de 25.000! Està clar que el tema té interès.



21. "Estadístiques", enquestes i experiments. Algunes idees

Passem ja a la segona capa. Parlarem dels conceptes clau, del mite de la grandària de la mostra i la grandària de la població, i d'alguns aspectes que no tenen res a veure amb cap càlcul, però que són molt importants.

Estimació de proporcions

Pregunta: Quina grandària de mostra necessito per que la informació sigui vàlida?

Resposta: Depèn

↓

- 1) Del marge d'error que s'estigui disposat a assumir
- 2) De la seguretat ("nivell de confiança") que es desitgi
- 3) Del valor que realment tingui la proporció que s'estimi
- 4) De la grandària de la població (però això es poc important si la població es gran)

22. "Estadístiques", enquestes i experiments. Algunes idees

Segurament una de les preguntes que més se'ns fa als que ens dediquem a l'estadística és aquesta de quina ha de ser la mida d'una mostra per que les conclusions siguin fiables i es puguin presentar a tot arreu (una altre pregunta freqüent es aquella de quina és la millor estratègia per que toqui alguna loteria).

La resposta a la pregunta –de la grandària de les mostres- és: “depèn”. I de què depèn?


- del marge d'error que s'estigui disposat a assumir.
- del nivell de confiança que es desitgi (ara veurem què són aquestes coses).
- del valor que realment tingui la proporció que s'estima. Això sembla estrany però és així. Si la proporció fos del 0 o del 100 % només caldria agafar una unitat; i a mesura que més s'acosta la proporció al 50% calen mostres més grans.
- i també depèn de la grandària de la població, tot i que això afecta menys del que sembla, com després veurem.

Interval de confiança per una proporció

Pregunta: ¿Quin percentatge de llars catalanes tenen ordinador amb connexió a internet?

Tema previ: Definicions clares
 Què és una llar?
 Pis d'estudiants?
 Apartament en la platja?...

Resultat: 42,8 %



Si tornéssim a fer l'estudi, tornaria a sortir el 42,8%?

23. "Estadístiques", enquestes i experiments. Algunes idees

Per exemple, podem fer una enquesta per estimar quin és el percentatge de llars catalanes que tenen ordinador amb connexió a Internet.

Tema previ: Què entenem per llar? Pis d'estudiants, apartament a la platja... És imprescindible tenir clars els termes que s'utilitzen, sinó la informació que es doni es pot prestar a confusions o males interpretacions.

Interval de confiança per una proporció

Pregunta: El 22% dels joves surten de nit sense la supervisió dels pares

Tema previ: Definicions clares

Resultat: El 22% dels joves surten de nit sense la supervisió dels pares

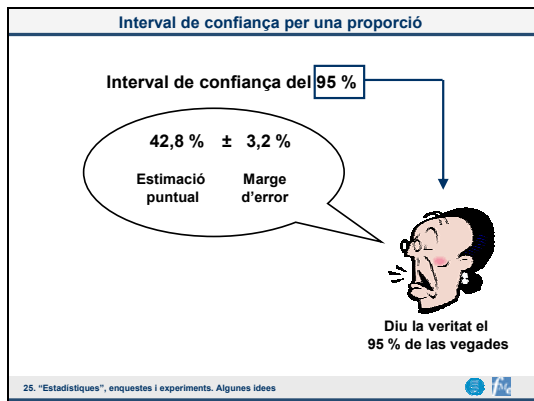
El 22% dels joves surten de nit sense la supervisió dels pares. Dit fins ara, es va portar a terme entre joves de 15 a 29 anys. Un 67% dels consultats surten normalment.

El Periódico, 30 gener 2006. Pàgina 06

24. "Estadístiques", enquestes i experiments. Algunes idees

Fem un incís sobre la claredat dels termes que s'utilitzen.

Mirem, per exemple, aquesta notícia. Diu que el 22% dels joves surt de nit sense supervisió dels pares. Sembla una dada preocupant, però si llegim el text veiem que quan parlen de joves s'estan referint a persones de 15 a 29 anys. És evident que per això de que el pares et supervisin quan surts no és el mateix tenir 15 que 29 anys.



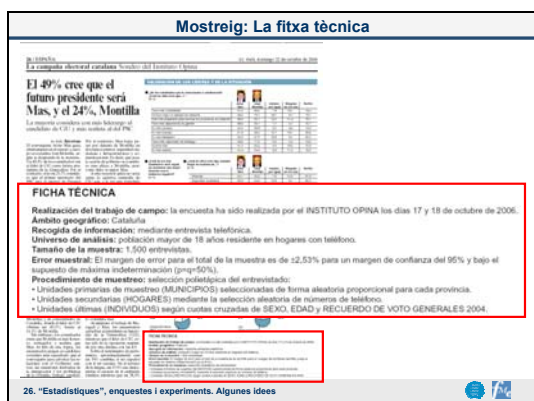
Ara suposem que ho hem fet tot bé, que la mostra és representativa, i que el resultat que ens ha sortit és que un 42,8% de la mostra té ordinador amb connexió a Internet. Si hem de fer una estimació d'un valor concret d'aquesta proporció en la població, amb la informació que tenim, el millor estimador seria 42,8%.

Però si tornéssim a fer l'estudi, tornaria a sortir el mateix valor? No. El més probable és que no, tampoc hauria de sortir un valor molt diferent, però el mateix no.

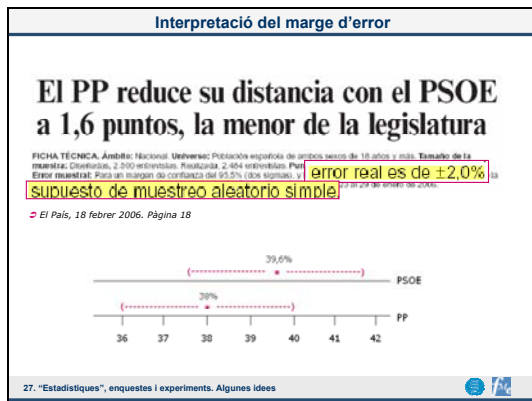
Per tant, els resultats de l'estudi es donen en forma d'interval. Direm que la nostra estimació està en l'interval 42,8 % ± 3,2%, sent aquest ± 3,2% el marge d'error.

Però, estem segurs de que el veritable valor d'aquesta proporció en la població està en l'interval que donem? Absolutament segurs no. Aquest interval ha estat calculat amb un determinat nivell de confiança. I això què vol dir? Si l'interval de confiança és del 95%, vol dir que s'ha calculat utilitzant un procediment que encerta el 95% de les vegades. És com si ens ho digués una persona que diu la veritat el 95% de les vegades.

I podríem calcular-lo amb una confiança del 99%? Sí, o més, però quant més gran sigui el nivell de confiança, també ho serà el marge d'error.



Veurem ara alguns retalls de diari. Ja és habitual en els diaris "seriosos" quan es presenten els resultats d'una enquesta acompanyar-los d'aquesta fitxa tècnica en la que expliquen com s'ha fet i quin és l'error de mostreig.



Però curiosament, sembla que les persones que redacten els titulars no entenen el significat d'aquests conceptes clau. Per exemple, aquí el titular destaca que el PP redueix la seva distància al PSOE a 1,6 punts... Però el text diu que l'error és del 2%, i això vol dir que, amb la confiança que ho han calculat, el valor real podria estar 2 punts amunt o 2 punts avall del que ha sortit. Que l'error sigui del 2% vol dir que si ho tornessin a fer, perfectament podria sortir 1 punt el PP per davant del PSOE. El titular correcte seria que no s'han detectat diferències significatives entre el PP i el PSOE.

Càlcul de la grandària de mostra per estimar una proporció

Interval de confiança $1-\alpha$ per P

$$p \pm z_{\alpha/2} \sqrt{\frac{N-n}{N} \frac{pq}{n-1}}$$

Marge d'error

$$E = z_{\alpha/2} \sqrt{\frac{N-n}{N} \frac{pq}{n-1}}$$

A partir de la fórmula anterior

$$E^2 = z_{\alpha/2}^2 \frac{N-n}{N} \frac{pq}{n-1}$$

En n-1 = n (aprox. raonable si n gran)

$$n = \frac{z_{\alpha/2}^2 pqN}{E^2 N + z_{\alpha/2}^2 pq}$$

$$n = \frac{z_{\alpha/2}^2 pqN}{E^2 N + z_{\alpha/2}^2 pq}$$

n: Grandària de la mostra
 $z_{\alpha/2}$: Lligat al nivell de confiança
 p: Proporció que es vol estimar
 q: 1-p
 N: Grandària de la població
 E: Marge d'error

28. "Estadísticas", enquestes i experiments. Algunes idees

A partir de l'expressió de l'interval de confiança és molt fàcil deduir una fórmula tancada que doni la grandària de la mostra en funció del nivell de confiança desitjat ($z_{\alpha/2}$), de les proporcions reals (p, q), la grandària de la població (N) i el marge d'error (E). Com que no coneixem les proporcions reals utilitzem els valors més desfavorables, que són p=q=0,5.

Càlcul de la grandària de mostra per estimar una proporció

29. "Estadísticas", enquestes i experiments. Algunes idees

I és molt fàcil crear-se en un full d'Excel un formulari que doni la grandària de la mostra en funció de les variables de les que depèn.

Grandària de la mostra i grandària de la població

Tamaño de muestra para estimación de proporciones con una confianza del 95%

Tamaño de la población	Márgen de error				
	±1%	±2%	3%	±4%	±5%
500	476	414	341	277	218
1.000	906	783	657	517	398
1.500	1.298	1.124	924	729	566
2.000	1.656	1.425	1.188	942	729
2.500	1.984	1.725	1.428	1.117	861
3.000	2.286	1.998	1.668	1.302	984
3.500	2.566	2.244	1.896	1.476	1.104
4.000	2.824	2.466	2.112	1.638	1.224
4.500	3.065	2.666	2.304	1.782	1.344
5.000	3.288	2.844	2.472	1.908	1.456
6.000	3.693	3.234	2.808	2.172	1.656
7.000	4.049	3.588	3.112	2.394	1.824
8.000	4.365	3.846	3.336	2.568	1.968
9.000	4.647	4.098	3.564	2.712	2.096
10.000	4.899	4.326	3.768	2.832	2.208
15.000	5.856	5.070	4.452	3.402	2.592
20.000	6.489	5.544	4.848	3.708	2.832
25.000	6.939	5.898	5.148	3.924	2.976
50.000	8.057	6.939	5.880	4.596	3.456
100.000	8.763	7.446	6.336	4.968	3.744
500.000	9.423	7.998	6.666	5.292	3.984
1.000.000	9.513	8.058	6.708	5.328	4.008
2.000.000	9.543	8.088	6.732	5.344	4.024
5.000.000	9.558	8.103	6.748	5.354	4.034
50.000.000	9.602	8.147	6.792	5.398	4.078

p=q=0,5. Valores redondeados por exceso

30. "Estadísticas", enquestes i experiments. Algunes idees

També amb l'Excel es pot construir directament una taula com aquesta que dóna les mides de mostra en funció de la mida de la població, el marge d'error i el nivell de confiança, en el cas més desfavorable ($p=q=0,5$).

Un dels aspectes més destacats d'aquesta taula és que la grandària de la mostra no augmenta proporcionalment amb la grandària de la població. Fixem-nos per exemple en un marge d'error del 4%: si la població és de 1.000 individus necessitem una mostra de 376, però si multipliquem per 10 la grandària de la població, la de la mostra no es multiplica ni per 2, i si a continuació multipliquem la població per 100, la mostra es queda pràcticament igual.

A moltes persones els hi costa entendre això. No entenen perquè es necessita la mateixa grandària de mostra per una població de 500.000 individus que per una de 50 milions.



Veurem un símil que ens permetrà veure de forma intuïtiva perquè això es així. Aquesta és l'olla en la que fem la sopa els dies de cada dia per la nostra família en la que som, posem per cas, 4 persones. Per veure si està bé de sal tenim una cullereta per tastar-la. I aquesta altra és l'olla que fem servir quan tenim convidats i ens ajuntem 10 o 12 persones a casa. La cullereta per tastar la sopa quan tenim convidats, com serà? Doncs serà la mateixa que la que fem servir els dies de cada dia. Encara que la "població" a tastar sigui més gran no cal que la grandària de la mostra augmenti. El que sí que cal, i és fonamental per l'èxit de l'operació, és remenar bé l'olla per assegurar que la mostra sigui representativa. És molt més important que l'olla estigui remenada i que, per tant, la mostra sigui representativa, que no pas que la cullera sigui més gran.

I el pecat de no remenar (no assegurar la representativitat de la mostra) no es perdona, com sembla que alguns pensen, augmentant la mida de la cullera (grandària de la mostra).

Muestreo: LA CLAVE

La "matemàtica" que hem vist
(càlcul del nivell de confiança, marge d'error, ...)
solament és vàlida si la mostra és


ALEATÒRIA

I aquest aspecte és fonamental tot i que
moltes vegades s'ignora

32. "Estadístiques", enquestes i experiments. Algunes idees

Cal recordar que és imprescindible que la mostra sigui aleatòria ("representativa", cal remenar l'olla); en cas contrari, tots aquests conceptes de marge d'error, nivell de confiança,... perden tot el seu significat.

Muestreo: Una encuesta sobre salarios



33. "Estadístiques", enquestes i experiments. Algunes idees

I també cal que la mostra es tregui de la població adequada. Si volem saber si el nou heroïnòman és jove o no, no podem preguntar només als joves...

Fent preguntes indiscretas...

Una experiència a la FME (Fòrum 2003):
Quin percentatge d'estudiants ha copiat alguna vegada als exàmens de l'FME?

L'enquestat tria una carta d'una baralla

- Si surt figura (sota, cavall, rei) respon a la pregunta innocent: El teu DNI és parell?
- Si no surt figura respon a la pregunta indiscreta: Has copiat mai als exàmens de la FME?

34. "Estadístiques", enquestes i experiments. Algunes idees

Quan es pregunta sobre el consum de drogues, vida sexual, o determinats comportaments mal vistos socialment o de caire personal, és molt fàcil que l'entrevistat enganyi perquè se senti violent confessant determinades coses a l'entrevistador (o entrevistadora). Però hi ha maneres d'aconseguir l'informació que permet estimar els paràmetres que interessin preservant l'intimitat de l'enquestat fins i tot davant l'entrevistador.

Us explicaré una experiència que varem fer aquí, a l'FME durant el Fòrum de l'any 2003, que és aquesta jornada en que les empreses munten estands i es posen en contacte els estudiants i les empreses. El nostre departament també munta un stand i organitza algunes activitats. L'any 2003 varem fer una enquesta (en un ambient bastant "festiu") per saber quanta gent havia copiat alguna vegada als exàmens de l'FME.

Si això es pregunta directament es corre el risc de que la persona no vulgui “confessar” i per assegurar que la seva resposta es mantenia en l’anonimat varem seguir el següent procediment. L’enquestat tria una carta d’una baralla que només ell mira i torna a posar-la a la baralla. El que ha de fer és:

- Si li ha sortit una figura ha de contestar a la pregunta. El teu DNI és parell?
- I si li surt una altra carta ha de contestar si ha copiat mai als exàmens de l’FME.

Si contesta “Sí”, l’entrevistador no pot saber si està contestant a la primera o a la segona pregunta.

Fent preguntes indiscretas...

$\alpha = 7/10$
 $1-\alpha = 3/10$

Has copiat mai?

DNI parell?

SI
NO

SI
NO

p
 $1-p$

$1/2$
 $1/2$

Probabilitat d'obtenir SI com a resposta: $\theta = \alpha p + \frac{1}{2}(1-\alpha)$

Estimador de θ : $\hat{\theta} = \frac{n_{SI}}{n}$ + Responen SI
 + Participen en l'enquesta

Estimador de p : $\hat{p} = \frac{\hat{\theta} - (1-\alpha)/2}{\alpha}$

35. "Estadístiques", enquestes i experiments. Algunes idees

L’anàlisi de les dades obtingudes d’aquesta forma és molt senzill. Aquí tenim les probabilitats que contesti a una pregunta o a una altra, i que contesti sí o no en cada cas. Nosaltres volem estimar el valor de p. La probabilitat d’obtenir sí com a resposta respon a aquesta expressió ($\theta = \alpha p + 1/2(1-\alpha)$), l’estimació de θ està clara i aïllar el valor de p és immediat.

Insisteixo en l’ambient festiu en què això es va fer. Un altre problema és que la pregunta no està massa clara. Què vol dir copiar? ...això no estava clar.

The screenshot shows a news article titled "Elecciones al Parlament 2006" from TTN. It features a bar chart with a rainbow color scheme representing the vote distribution among different political parties. Below the chart is a table with columns for "PROVINCIA", "CIUTADANOS", "DIFER", and "INDIVIDUALS". The rows are labeled "Barcelona", "Girona", and "Lleida".

Acabaré aquesta part de les enquestes fent uns comentaris sobre els sondejos electorals.

Ja sabeu que endevinar qui guanyarà les eleccions o com quedarà el parlament és complicat (i encara més saber qui serà President). Les raons són variades, la gent no sap a qui votarà, o no contesta, o enganya, o després canvia d’opinió. A més, una cosa és estimar proporcions i una altra estimar escons. Petites variacions en la proporció de vots (que no es poden discriminar amb les tècniques estadístiques que s’utilitzen) poden afectar a que un escó sigui atribuït a un o altre partit...

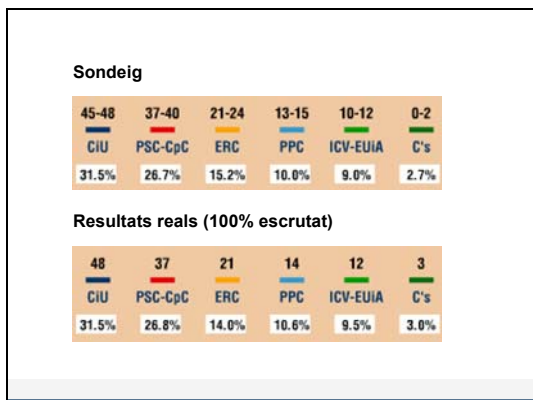
I hi ha una altra raó per la que el sondejos electorals

tenen mala fama. I és que quan s'equivoquen se'n parla molt. Fa gràcia, són notícia. En canvi, quan l'endevinen no té tant interès i no se'n parla tant.

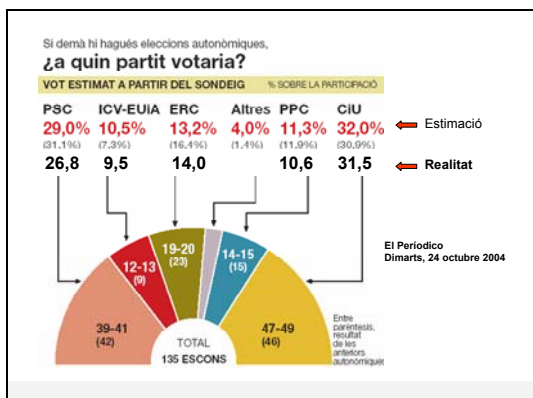
Aquí tenim la reproducció de la pàgina web de TV3 (o de Catalunya Informació) la nit de les eleccions, a les 20:38 amb el resultat del sondeig que s'havia fet a peu d'urna i que ja van anunciar a les 20 en punt.



I aquesta és la mateixa pàgina actualitzada a les 00:20 amb els resultats definitius.



S'hi assemblen? Doncs sí, i molt...



Aquestes són les dades d'un sondeig que es va publicar una setmana abans de les eleccions. No és exacte, però no està malament... El sondejos no sempre s'equivoquen!

Contingut
Primera capa: Anàlisi exploratòria de dades Mesurar la variabilitat Percentatges Gràfics
Segona capa: Mostreig Conceptes clau Grandària de mostra Males interpretacions Més enllà dels càlculs...
Tercera capa: Experiments Experimentació i estadística? Alguns exemples
<small>40. "Estadístiques", enquestes i experiments. Algunes idees</small>

I entrem en l'última capa per parlar de l'experimentació i el disseny d'experiments, a la que dedicaré menys temps que a les anteriors, però sobre la que m'agradarà fer uns comentaris per donar una visió més completa dels camps d'actuació de l'estadística.


Disseny d'experiments	
El raonament estadístic i l'avenç de la ciència	
 <p>Ronald A. Fisher (1890-1962)</p>	
<small>41. "Estadístiques", enquestes i experiments. Algunes idees</small>	

A principis del segle XX es va produir un canvi important en la metodologia d'investigació i d'adquisició de nous coneixements a partir de la recollida i l'anàlisi de dades. I un dels protagonistes més destacats d'aquesta revolució va ser Ronald Fisher, tot i que segurament és un personatge del que se'n parla poc.

En un llibre que Fisher va escriure sobre disseny d'experiments explica l'anècdota d'una senyora que afirmava ser capaç de distingir quan en una tassa de te amb llet, s'havia posat primer el te o la llet, i explica com va organitzar una prova per esbrinar si aquesta senyora era realment capaç de distingir les 2 situacions.

El plantejament de la prova i la forma de raonar ara ens semblen bastant evidents, però en aquella època, sobre els anys 20 del segle XX eren idees originals. "The lady tasting tea" és un llibre molt interessant i entretingut, que utilitza aquesta anècdota com a fil conductor per explicar com l'estadística va revolucionar la ciència al segle XX.

El cas de la tastadora de te



$p = 0,5$

$p = \frac{1}{C_2^4} = \frac{1}{6} = 0,17$

$p = \frac{1}{C_8^4} = \frac{1}{70} = 0,0143$

42. "Estadístiques", enquestes i experiments. Algunes idees

Mirem una mica el raonament. Si a aquesta persona li donem dues tasses i li diem que en una s'ha posat primer el te i en l'altra primer la llet, i les ha d'identificar, té una probabilitat del 50% d'encertar-ho a l'atzar i, per tant, que ho encerti no és cap garantia de que realment sigui capaç de diferenciar les dues situacions.

Si li donem 4 tasses ja és més difícil que ho endevini per atzar. Es poden fer 6 parelles diferents... i la probabilitat de que ho endevini a l'atzar és del 17%. Potser encara és massa gran i si ho endevinés tampoc estariem segurs de que realment és capaç de distingir les dues situacions. Si li donem a tastar 8, la probabilitat de que ho endevini per atzar és només de l'1,4%. Si estem disposats a córrer un risc de fins al 5%, aquest podria ser un bon disseny.

Es pot disminuir la probabilitat d'endevinar-ho a l'atzar a base de que la tastadora no sàpiga quantes tasses hi ha de cada tipus... i hi ha diferents procediments que tenen els seus avantatges i els seus inconvenients... però la idea clau és que s'ha dissenyat una prova per contrastar una hipòtesi a partir de les dades obtingudes i amb una probabilitat d'error coneguda i que es pot fixar segons convingui.

Tast d'aigües




Rigor metodològic
Doble cec

43. "Estadístiques", enquestes i experiments. Algunes idees


A les classes de disseny d'experiments fem unes proves similars però no amb tes sinó amb dos tipus d'aigua (Font Vella i de les fonts de la facultat). Cal adonar-se que no és només un problema de càlcul de probabilitats. És necessari dissenyar la prova amb el rigor adequat, i de manera que el que serveix l'aigua i anota els resultats tampoc sàpiga de quin origen és l'aigua de cada ampolla (doble cec), perquè si ho sabés podria ser que a través d'un gest, d'una mirada o d'un comentari donés pistes al tastador sobre si s'està equivocant, o no...

Proves similars

Serveix la cullereta per que el cava no s'esbravi?



Poden identificar quin meló és el millor?



Cóm dissenyar aquestes proves?
Quins aspectes s'haurien de tenir en compte?

44. "Estadístiques", enquestes i experiments. Algunes idees

Proves similars es poden fer per contrastar la veracitat de llegendes urbanes com la de que si es deixa un cullereta en una ampolla de cava es conserva millor que si es deixa oberta.

O una altra, que crec que és similar, és contrastar si realment aquestes persones que tocant un meló i observant no sé quins detalls, poden saber si estan bons o seria millor menjar-los demà... És a dir, si realment tenen raó o ho endevinen les mateixes vegades que si escollissin a l'atzar.

L'aspirina i les flors del gerro




45. "Estadístiques", enquestes i experiments. Algunes idees

Però hi ha molts tipus d'experiments. Un típic és provar si realment una aspirina en l'aigua d'un gerro allarga la vida de les flors. Cal ser meticulós i ordenat en el disseny de l'experiment.

Aspectes a considerar

Flor	Tipus Estat inicial Quantitat	Gerro	Forma Material Transparència
Tija	Longitud Tocar al fons Tall	Ambient	Temperatura Lluminositat Hores de llum
Aigua	Origen Composició Quantitat		
Aspirina	Forma Tipus		

46. "Estadístiques", enquestes i experiments. Algunes idees


Cal identificar tots els aspectes que cal tenir en compte, i ser conscient de l'àmbit en que seran vàlids els resultats de l'experiment...

Mesura de la resposta

Bona



Dolenta



47. "Estadístiques", enquestes i experiments. Algunes idees

Cal definir d'una forma clara i objectiva quan una flor està pansida per poder valorar quants dies ha durat.

Registre de les respostes

Facha	Serons			
	1	2	3	4
D 1987				
D 1988				
L 1989				
L 1990				
M 1991				
M 1992				
M 1993				
J 1994				
J 1995				
V 1996				
S 1997				
S 1998				
D 1999				
L 2000				
L 2001				
M 2002				
M 2003				
M 2004				
M 2005				
M 2006				
M 2007				
M 2008				
M 2009				
M 2010				
M 2011				
M 2012				
M 2013				
M 2014				
M 2015				
M 2016				
M 2017				
M 2018				
M 2019				
M 2020				
M 2021				

48. "Estadístiques", enquestes i experiments. Algunes idees

Cal, a més, dissenyar un procediment per recollir les dades d'una forma clara i ordenada...

Comparació d'helicòpters







49. "Estadístiques", enquestes i experiments. Algunes idees

També es poden dissenyar experiments per comparar els temps de caiguda de diferents tipus d'helicòpter de paper. Quants helicòpters fan falta de cada tipus? Metodològicament, ¿és igual construir un sol helicòpter d'un cert tipus i deixar-lo caure 10 vegades, per exemple, que construir 10 helicòpters diferents i deixar-los caure una vegada cada un?

La meua experiència és que aquests pràctiques són divertides i molt formatives.

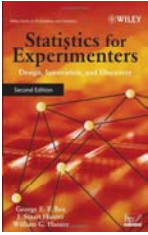
Més experiments...

Duren més les piles cares que les barates

Serveixen les bosses de plàstic plenes d'aigua per espantar les mosques?

És veritat que el detergent X renta més blanc?

És veritat que el rentavaixelles Y rendeix més?



50. "Estadístiques", enquestes i experiments. Algunes idees

I es poden fer experiments per comprovar si les piles cares duren més que les piles barates. Tots sabem que les piles de determinades marques són bastant més cares que les de marques blanques (com les que es venen al Lidl), però no és suficient provar una de cada marca perquè les d'un cert tipus no duren totes el mateix. I com mesurem el temps que dura una pila en una llanterna? No estarem tota l'estona mirant fins que s'acabi. Pot ser dura molt en una llanterna però molt poc en un aparell mp3 o en una càmera de fotos...

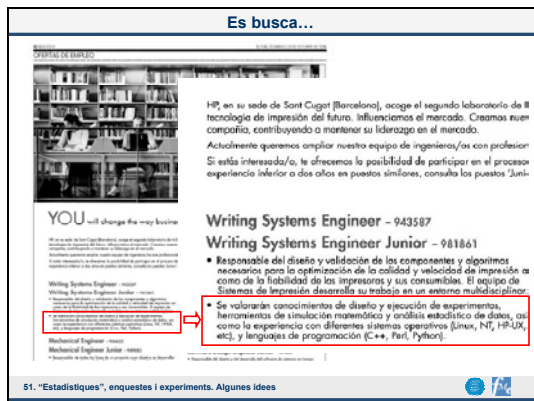
Altres preguntes que ens podem plantejar són:

- Serveixen les bosses de plàstic plenes d'aigua per espantar les mosques?
- És veritat que el detergent X renta més blanc?

Com faríem les proves? Com analitzaríem els resultats?

En totes aquestes situacions en les que la variabilitat està present (no sempre els resultats són els mateixos encara que provem sempre el mateix producte), i ho està quasi sempre, cal utilitzar l'estadística.

El llibre d'estadística que més m'agrada és aquest ("Statistics for Experimenters"), que és molt conegut, i que parla de l'estadística en l'experimentació, i de disseny, innovació i descobriments... paraules molt actuals i que crec que tenen molt futur.



Per acabar us volia ensenyar un retall de diari que no és una pífia estadística sinó una oferta de feina. La vaig veure fa un parell de setmanes quan començava a pensar en el que us explicaria.

És un anunci singular, que ocupa tota una plana del diari de la qual aproximadament 1/3 és una fotografia. És per la planta de Sant Cugat de Hewlett Packard, el lloc on es dissenyen les impressores de gran format que es fabriquen i es comercialitzen arreu del món, i segurament un dels centres del nostre país on més es practica això de la recerca i la innovació.

Busquen diferents tipus de tècnic. I concretament pel primer, “expert en sistemes d’escriptura” se li valoraran coneixements de disseny i execució d’experiments, eines de simulació matemàtica, anàlisi estadística de dades i informàtica. Crec que aquests disciplines tenen molt a veure amb la competitivitat i amb el progrés i són cada vegada més importants en la nostra societat.

Moltes gràcies.

Crèdits

Moltes de les idees que apareixen en aquesta xerrada les he anat recollint dels meus companys professors d’estadística, i també dels meus estudiants. La idea d’utilitzar exemples de marciàns per discutir el concepte de desviació tipus li vaig sentir a Xavier Tort-Martorell. En Xavier també em va suggerir incloure la foto de “l’histograma viu” en aquesta xerrada. La majoria de notícies dels diaris les he tret del projecte final de carrera que van realitzar Sara Fontdecaba i Maria Montón sobre l’ús de l’estadística en la premsa. L’exemple de l’olla i la cullereta és del nostre amic Roberto Behar, professor d’estadística a la Universidad del Valle, a Colòmbia, i un apassionat de la divulgació de l’estadística. L’enquesta sobre quants estudiants copien a l’FME la va organitzar Pedro Delicado, i d’un petit informe que va escriure he tret els comentaris per aquesta presentació. En Lluís Marco, coautor dels pòsters de les pífies, m’ha ajudat en la preparació de tot el material i, en concret, m’ha donat els resultats dels sondeigs de les últimes eleccions. Ignasi Solé i Lourdes Roderó em van fer diversos suggeriments que m’han servit per explicar millor algunes coses. Maria Montón ha preparat aquesta versió escrita.

Pere Grima, 22 de novembre de 2006



El text i les imatges d’aquest document estan subjectes a una llicència Reconeixement-NoComercial-SenseObraDerivada 2.5 de Creative Commons. Podeu copiar-los, distribuir-los i comunicar-los públicament sempre que citeu la font, no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/2.5/es/deed.ca>.