Campionamento

Come circoscrivere -grazie al calcolo delle probabilità- il numero di unità e come realizzare la loro scelta.

C'è un duplice aspetto:

Procedura di selezione

Insieme di regole ed operazioni con cui si realizza la scelta delle unità;

Procedura di stima

Calcolo delle statistiche e loro uso come valori presunti dei parametri incogniti.

Per ora ci interessa solo il primo

Campionamento/2

Se si considerano tutte le unità, la probabilità non svolge alcun ruolo.

Se non è possibile effettuare un'indagine completa ci saranno unità effettivamente esaminate ed altre no.

La scelta può avvenire assegnando ad ogni unità una certa probabilità di comparire nell'indagine:

Il risultato è che ci si trova di fronte a dei dati che sono quelli, ma avrebbero potuto essere altri. Cosa si può dire allora sui risultati ottenuti?

Campionamento/3

Operare con dei campioni significa agire con informazioni incomplete

Da un lato c'è la POPOLAZIONE: un collettivo di oggetti, persone, valori per la quale si cerca di conosceere un fatto

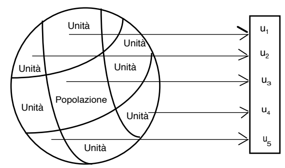
Dall'altro c'è il CAMPIONE: una parte della popolazione che viene esaminata perché possa fornire informazioni sulla popolazione

Dobbiamo perciò essere preparati ad un errore dovuto alla mancanza di dati relativi alle unità non campionate.

C'è un modo per minimizzare lo scarto tra valori campionari e valori nella popolazione?

Popolazione teorica ed effettiva

In ogni rilevazione di dati è necessario un sistema di riconoscimento delle unità



Non sempre è facile fare la lista delle unità: popolazioni elusive od incerte in cui non sappiamo chi sono e quante sono le unità

Dobbiamo quindi distinguere tra

- POPOLAZIONE TEORICA: quella che vorremmo esaminare
- POPOLAZIONE EFFETTIVA: unità effettivamente raggiungibili

La *frame* o lista

Tra TEORICA ed EFFETTIVA si inserisce la *frame* o lista cioè un sistema di codici che identificano le unità

La lista è una sovrastruttura imposta alla popolazione.

Per essere utile deve risultare:

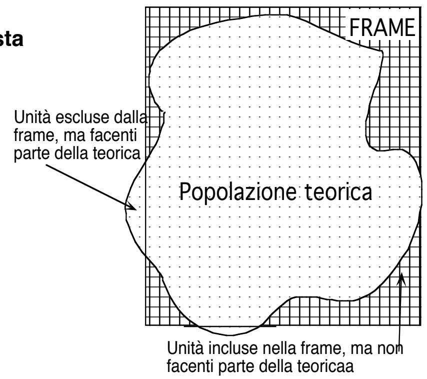
ESATTA

AGGIORANTA

COMPLETA

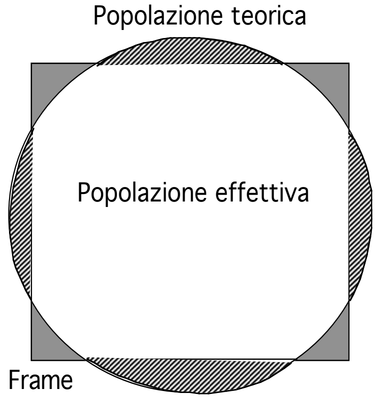
DOCUMENTATA

con regole note e disponibili



Le ragioni della scelta di una certa popolazione effettiva per una data popolazione teorica devono essere capillari ed accurate.

La *frame/2*



La popolazione effettiva è data dalle unità della popolazione teorica incluse nella *frame*.

La copertura assicurata dalla popolazione effettiva è soddisfacente se rende superflua la ricerca sulla teorica.

Una frame è perfetta se tutte le unità della popolazione vi sono elencate una ed una sola volta.

Errori di frame

UNDERCOVERAGE **Unità della popolazione teorica non viste dalle *frame***

OVERCOVERAGE **Unità viste dalle *frame* ma non facenti parte della teorica (include i doppioni)**

NONCOVERAGE **Unità della popolazione teorica viste dalle *frame* ma dalle quali non è stato possibile avere alcun dato**

NON RESPONSE **Per quanto si faccia non si riescono ad acquisire, dalla unità, le informazioni necessarie**

Tali errori impediscono un corretto collegamento (selection bias) tra teorica ed effettiva molto pericoloso per le indagini.

Errori non campionari

DISTORSIONI NELLA SCELTA DELLE UNITA'

Il meccanismo di estrazione delle unità agisce solo su alcune parti e ne esclude SISTEMATICAMENTE altre (undercoverage e over coverage)

ESEMPIO:

i sondaggi telefonici: in dipendenza dell'orario in cui si telefona si raggiungono unità diverse. Persone che non hanno telefono o il cui telefono non è in elenco non potranno mai essere raggiunte.

DISTORSIONI NELLA RILEVAZIONE DELLE UNITA'

Non sempre è possibile garantire l'accuratezza delle misurazioni ovvero non sempre le unità sono disposte a farsi rilevare o a rispondere con sincerità

ESEMPIO:

Rilevazione degli affiliati a "Cosa Nostra"

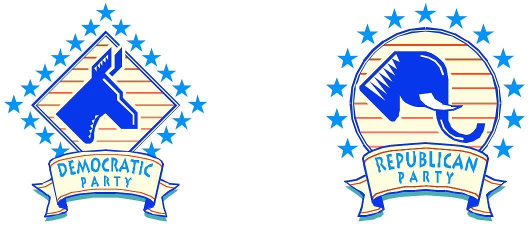
Il noto caso del Literary Digest

Nel 1936 tale rivista attivò un sondaggio postale su dieci milioni di votanti scelti da elenchi telefonici e registri di possessori di auto.

Lo scopo era di prevedere il risultato delle elezioni presidenziali: Roosvelt (democratici-progressisti) e Landon (repubblicani-conservatori).

Si ottennero 2.4 milioni di riposte: il 57% avrebbe votato Landon ed il 38.5% Roosvelt. Vinse però Roosvelt con il 63%. Gran parte del fiasco è da attribuire ad una scelta inadeguata della lista.

Perché?



L'ampiezza del campione

il numero di unità “n” considerate nel campione è detto AMPIEZZA del campione. L'ampiezza della popolazione è indicata con "N"

$Frazione\ \ di\ \ campionamento: \ \frac{n}{N}$ Frazione di unità della popolazione "rappresentata"nel campione

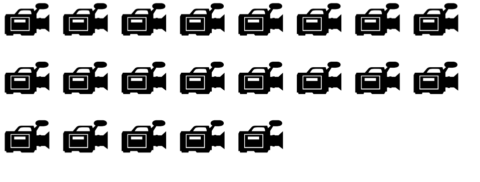
$Intervallo\ \ di\ \ campionamento: \ \frac{N}{n}$ Ogni quante unità della popolazione si sceglie una unità del campione.

Anche, fattore di proporzionalità.

Se la popolazione è infinita questi due rapporti perdono di significato.

Variabilità della popolazione

Se le unità fossero uguali basterebbe un campione di ampiezza n=1



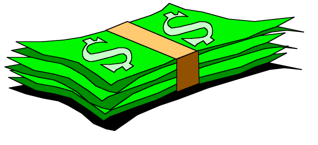
Se le modalità sono due o più saremo certi che entrambe siano nel campione solo se n=N.

Ma la variabilità non è nota. Spesso è uno degli scopi di una ricerca.

Determinare "n"

E’ un elemento fondamentale del piano di campionamento. Sulla scelta incidono...

- Obiettivo dell'indagine
- Variabilità attesa nella popolazione (controllo errori campionari)
- Costo dell'acquisizione



- Controllo errori non campionari

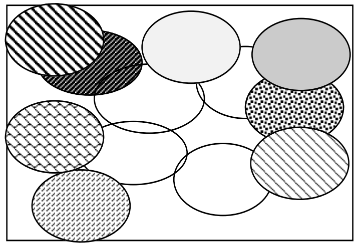
La determinazione di "n" è molto complessa e verrà ripresa in un corso successivo

L'universo dei campioni

Da una data popolazione è possibile selezionare un certo numero di campioni.

Fissata l'ampiezza campionaria "n" definiamo UNIVERSO DEI CAMPIONI (di ampiezza "n") l'insieme di tutti campioni di tale ampiezza che possono essere ottenuti da una popolazione "P"

$$T_n(P) = \{C_1, C_2, \ldots, C_i, \ldots\}$$



L'universo dei campioni può anche essere considerato una POPOLAZIONE le cui unità sono i campioni di ampiezza "n"

Naturalmente si tratta di una popolazione astratta che sarà censita quasi sempre in una sola unità

Cardinalità dell'universo dei campioni

Dipende …

- **Dalla possibilità di ripetere o no la stessa unità**
- **Se rileva o no l'ordine di comparizione nel campione.**

Se non c'è reimmissione ed i campioni sono considerati uguali purché formati dalle stesse unità allora la cardinalità è il coefficiente binomiale:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N*(N-1)*(N-2)*\ldots*(N-n+1)}{n!}$$

Esempio:
Ad un test sull'impatto visivo di un poster 6x3 metri sono stati invitati N=50 automobilisti che hanno dato la loro opinione. Di questi, n=7 dovrebbero essere sottoposti ad un altro test sulla leggibilità delle scritte.

Le scelte possibili sono:

$$\binom{50}{7} = \frac{50!}{7!43!} = 99'884'400$$

Cardinalità2

Se le unità possono ripetersi fermo restando che l'estrazione è senza reimmissione e che l'ordine non conta, la cardinalità è:

$$\binom{N+n-1}{n} = \frac{N*(N+1)*\ldots*(N+n-1)}{n!}$$

Esempio:
Si analizzano le N=70 sentenze emesse da un collegio giudicante con un'intervista di n=6 condannati. La presenza di recidivi può provocare la ripetizione delle unità.

I campioni possibili sono

$$\binom{70+6-1}{6} = \frac{70*71*\ldots*75}{6!} = 201'359'550$$

Cardinalità/3

Se è NON consentita la reimmissione e l'ordine diverso rende diversi due campioni con uguali unità allora i campioni possibili sono

$$\frac{N!}{(N-n)!}$$

Esempio:
La famosa scienziata ha intuito che un trattamento di 5 farmaci scelti tra 20 principi attivi e somministrati nel giusto ordine può curare una fastidiosa patologia.

Scegliendo a caso i principi attivi quanti campioni sono possibili?

$$\frac{20!}{15!} = 20*19*18*17*16 = 1'860'480$$

Cardinalità/4

Se è consentita la reimmissione e l'ordine diverso rende diversi due campioni con uguali unità allora l'universo dei campioni ha un numero di elementi pari a:

$$N^n$$

Nei due esempi precedenti avremmo:

$$50^7 = 781'250'000'000; \quad 70^6 = 117'649'000'000$$

il numero di elementi dell'universo dei campioni è quasi sempre troppo elevato -anche con i supercalcolatori- perché valga la pena di studiare il comportamento delle statistiche su tutti.

La rappresentatività

il campione deve essere RAPPRESENTATIVO della popolazione da cui è estratto, cioè assicurare che i risultati qui ottenuti si estendano a tutta la popolazione. Almeno in relazione alla caratteristica in esame

I giocatori di una squadra di basket non sono rappresentativi della popolazione per l'altezza, ma potrebbero esserlo per capacità di apprendimento.

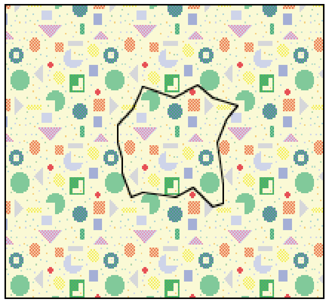
Solo la rilevazione totale (senza errori non campionari) è certamente rappresentativa ovvero la selezione di un numero qualsiasi di unità da una popolazione invariabile.

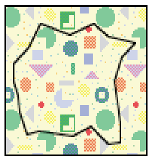
In entrambi i casi il campione è inutile

La rappresentatività/2

La figlia vuole un vestito con il medesimo disegno di quello della madre. Che campione si dovrà portare al negozio di stoffe?

POPOLAZIONE





Campione

Il campione deve essere abbastanza piccolo per evitare di impacchettare l'intero vestito, ma deve anche essere abbastanza grande da includere il motivo ricorrente della stoffa.

La formazione del campione

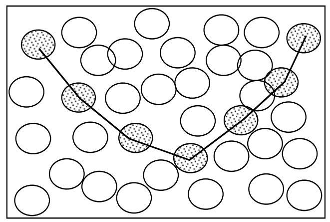
L'efficacia di un'indagine parziale dipende dalla capacità di mimare e miniare la rilevazione totale cui si sostituisce cioè la sua RAPPRESENTATIVITA'

La rappresentatività non riguarda il campione in quanto tale, ma il metodo di scelta delle sue unità

Ci sono tanti campioni possibili

Alcuni sono rappresentativi cioè prossimi alla popolazione per gli aspetti che ci interessano

Altri non sono rappresentativi e danno una idea distorta delle variabili di interesse.



Il metodo di campionamento (sample design) si concretizza nella scelta del campione con una sequenza di scelte che tende a produrre campioni rappresentativi.

il campione casuale

Un tipico esperimento che rientra nel modello delle urne è la scelta casuale di "n" unità da una popolazione finita di "N".

L'evento elementare è la n-tupla di interi $C_i=(i_1,i_2,\ldots,i_n)$ corrispondenti alle posizioni occupate in una lista univoca ed esaustiva delle unità della popolazione;

1	2	$\ldots$	$i-1$	i	$i+1$	$\ldots$	$n-2$	$n-1$	n
u	u		u	u	u		u	u	u

Lista o frame

il campione casuale semplice presuppone che tutte le unità siano scelte con uguale probabilità

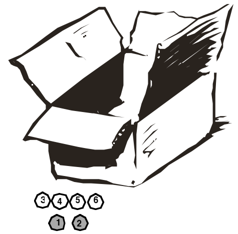
Casualità e campionamento

Se la popolazione è grande si usano tecniche informatiche per scegliere le unità da includere nel campione

1	2	$\ldots$	$i-1$	i	$i+1$	$\ldots$	$n-2$	$n-1$	n
u	u		u	u	u		u	u	u

Per ampiezze più piccole basta una scatola con delle biglie di uguale volume, peso, superficie, temperatura, porosità, colore, lucidatura, etc.

Prima di ogni estrazione la scatola (opaca) è ben scossa con moto sussultorio e ondulatorio che renda impossibile predeterminare la posizione di una particolare biglia



Campione casuale semplice con reimmissione

Scelta di n=3 famiglie in una lista di N=100. Supponiamo che, dopo ogni estrazione, la biglia sia rimessa nell'urna che poi è adeguatamente scossa

La procedura descritta assicura che:

Ognuna delle N famiglie della popolazione ha la stessa probabilità di comparire in una qualsiasi delle n posizioni del campione:

$$P(Fam_i \ \ sia \ \ in \ \ posizione \ \ j)=\frac{100*100}{100*100*100}=\frac{1}{100} \qquad \frac{casi \ \ favorevoli}{casi \ \ possibili}$$

Ne consegue che ogni gruppo di n famiglie ha la stessa probabilità di costituire il campione:

$$P\left(Fam_{i_1}, \ Fam_{i_2}, \ Fam_{i_3}\right)=\frac{1}{100*100*100}$$

C.C.S. in blocco

Dopo ogni estrazione, la biglia resta fuori dell'urna che viene comunque adeguatamente rimischiata.

La probabilità che la famiglia "i" compaia al 1° posto del campione è 1/(99*98)

Fam_i	2°	3°

il 1° posto e bloccato, il 2° può essere occupato dalle 99 restanti famiglie ed il 3° in 98 modi diversi dato che due famiglie impegnano già i primi due posti.

La stessa cosa succede per tutte le altre posizioni

1°	Fam_i	3°
99	1	98

1°	2°	Fam_i
99	98	1

Ognuna delle N famiglie della popolazione ha la stessa probabilità di comparire in una qualsiasi delle n posizioni del campione:

$$P(Fam_i \ \ sia \ \ in \ \ posizione \ \ j)=\frac{99*98}{100*99*98}=\frac{1}{100}$$

C.C.S. in blocco/2

Scelta la prima famiglia su N a far parte del campione, la seconda è scelta su 99, la terza su 98.

Qualunque famiglia può essere la prima, la seconda o la terza.

Ne consegue che:

Ogni gruppo di n famiglie ha la stessa probabilità di costituire il campione:

$$P\left(Fam_{i_1},\ Fam_{i_2},\ Fam_{i_3}\right)=\frac{1}{100*99*98}$$

Cardinalità dell'universo dei campioni

Proprietà del campione casuale

Il campione casuale ha le proprietà seguenti:

1) Ogni unità può comparire in una qualsiasi delle posizioni del campione.

2) Se le unità sono equiprobabili, ogni gruppo di unità ha la stessa probabilità di formare il campione

3) Se le unità sono scelte con reimmisione le singole scelte sono indipendenti

4) Se il campione è piccolo rispetto alla popolazione la differenza dovuta alla reimmissione/non reimmissione diventa trascurabile

Casualità e campionamento/2

E' casuale il meccanismo di scelta delle unità e non il campione ottenuto.

Perché scegliendo a caso dalla popolazione si può ottenere un campione rappresentativo?

Una popolazione è formata da tre tipi di unità: A, B, C di cui è nota la proporzione nella popolazione: p(A)=50%, p(B)=30%, p(C)=20%.

All'aumentare **dell' ampiezza del campione, il meccanismo casuale di scelta porta a campioni che riproducono la popolazione.**

Ampiezza	A	B	C
n=10	0.6	0.2	0.2
n=100	0.51	0.29	0.2
n=1000	0.512	0.293	0.195
n=10000	0.5017	0.2983	0.2
n=100000	0.50047	0.302	0.19573
n=1000000	0.500482	0.301929	0.197589
Popolazione	0.5	0.3	0.2

Questo è il postulato empirico del caso

Casualità e campionamento/3

L'esperimento, come tanti altri dello stesso genere, mostra che in un campione casuale abbastanza grande, le unità sono guidate dalla sorte a comparire nelle medesime proporzioni con cui sono presenti nella popolazione.

Il singolo campione può avere una conformazione più o meno simile a quello della popolazione, ma non è dato sapere in che misura (a meno che non si tratti di una simulazione in cui la popolazione sia nota).

Ciò che tranquillizza gli utilizzatori della Statistica è che un sorteggio corretto delle unità e per ampiezze campionarie elevate tenderà a riprodurre le caratteristiche della popolazione.

Errori campionari

L'uso del campione introduce un errore dovuto alle differenze tra risultati nel campione e risultati POTENZIALI ottenibili dall'esame di tutta la popolazione.

Gli errori sono dovuti a fluttuazioni in parte attribuibili alla naturale variabilità campionaria: i dati sono quelli, ma potevano essere altri

ESEMPIO
Vogliamo conoscere il totale dei valori della tabella (popolazione).

Si sceglie una riga o una colonna di cinque numeri (campione)

7	13	5	5	10
2	8	5	4	1
6	10	11	1	12
1	7	8	4	8
2	3	3	1	3

Riga	Stima	Errore camp.	Colonna	Stima	Errore camp.
40	200	60	18	90	-50
20	100	-40	41	205	65
40	200	60	32	160	20
28	140	0	15	75	-65
12	60	-80	34	170	30

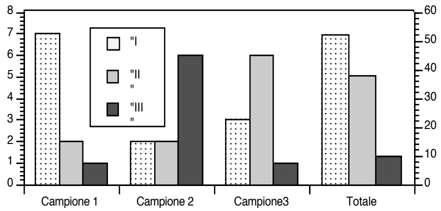
Solo se si sceglie la 4ª riga non c'è errore campionario

Errori campionari/2

Qualunque sia la conclusione raggiunta a mezzo del campione essa è venata da incertezza.

il suo successo può corroborare la validità della procedura per il passato, ma ben poco può aggiungere sulla conoscenza del suo comportamento futuro.

ESEMPIO:
una variabile può assumere tre soli valori: 1,2,3. Per stimare la sua distribuzione di frequenza: per campione si sceglie: prima colonna, ultima colonna e terza riga.



Popolazione dei valori

	1	2	3	4	5	6	7	8	8	9	10
1	1	1	1	3	1	1	2	3	1	1	2
2	1	1	1	3	2	3	1	1	2	2	1
3	2	2	2	2	2	2	2	1	1	1	3
4	3	3	2	1	3	2	1	2	1	2	3
5	3	1	3	1	1	2	1	1	1	3	3
6	1	1	1	2	2	1	2	2	1	1	2
7	1	3	2	3	1	1	2	2	2	1	1
8	1	1	1	1	2	2	1	1	2	1	3
9	1	2	3	1	2	2	1	3	3	2	3
10	1	2	2	1	2	1	1	1	3	2	3

Per ciascuno dei campioni si prenderà una decisione sbagliata

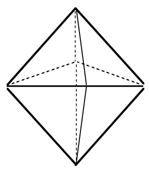
Sorteggio delle unità

il modello della pura sorte per scegliere le unità può essere simulato in molti modi: monete, dadi.

Ogni processo fisico che nel suo funzionamento segua lo schema del sorteggio tra unità identiche può servire a formare il campione

ESEMPIO: l'ottaedro ha 8 facce uguali a forma di triangolo. Se fatto rotolare su di una superficie liscia finirà col poggiarsi su di una faccia. Se è ben costruito risulta imprevedibile la faccia su cui si poggerà

Accostando le uscite si ottiene il numero casuale in base ottale: 7201 convertibile in base decimale: $7*8^3+2*8^2+0*8+1= 3713$ o tra zero ed uno dividendo per 4095



I giochi di sorte sono emblematici: le uscite sono casuali se nessun giocatore -per quanto furbo- trova una regola che gli consenta di scommettere meglio che alla pari.

Estrazione con i numeri casuali

Nella pratica che precedette l'home computing il processo di estrazione dall'urna era spesso simulato mediante dei numeri casuali disponibili in forma di tabelle

Le tavole dei numeri casuali sono delle raccolte di numeri da 0 a 9 variamente raggruppate e caratterizzate dall'assenza di una legge di successione o di ordinamento.

La costruzione è fatta in modo che le cifre da "0" a "9" hanno ciascuna frequenza 1/10 di ripetersi, le coppie da "00" a "99"frequenza 1/100, le terne "000"…"999" hanno 1/1000 ecc.

Esempio di tabella di numeri casuali

	1	2	3	4
1	53 74 23 99 67 63 38 06 86 54 35 30 58 21 46 63 43 36 82 69 98 25 37 55 26	61 32 28 69 84 99 00 65 26 94 06 72 17 10 94 65 51 18 37 88 01 91 82 81 46	94 62 67 86 24 02 82 90 23 07 25 21 31 75 96 61 38 44 12 45 74 71 12 94 97	98 33 41 19 9S 79 62 67 80 60 49 28 24 00 49 32 92 85 88 65 24 02 71 37 07
2	02 63 21 17 69 64 55 22 21 82 85 07 26 13 89 58 54 16 24 15 34 85 27 84 87	71 50 80 89 56 48 22 28 06 00 01 10 07 82 04 51 54 44 82 00 61 48 64 56 26	38 15 70 11 48 61 54 13 43 91 59 63 69 36 03 62 61 65 04 69 90 18 48 13 26	43 40 45 86 98 82 78 12 23 29 69 11 15 83 80 38 18 65 18 97 37 70 15 42 57
3	03 92 18 27 46 62 95 30 27 59 08 45 93 15 22 07 08 55 18 40 01 85 89 95 66	57 99 16 96 56 37 75 41 66 48 60 21 75 46 91 45 44 75 13 90 51 10 19 34 88	30 33 72 85 22 86 97 80 61 45 98 77 27 85 42 24 94 96 61 02 15 84 97 19 75	84 64 38 56 98 23 53 04 01 63 28 88 61 08 84 57 55 66 83 15 12 76 39 43 78
4	72 84 71 14 35 88 78 28 16 84 45 17 75 65 57 96 76 28 12 54 43 31 67 72 30	19 11 58 49 26 13 52 53 94 53 28 40 19 72 12 22 01 11 94 25 24 02 94 08 63	50 11 17 17 76 75 45 69 30 96 25 12 74 75 67 71 96 16 16 88 38 32 36 66 02	86 31 57 20 18 73 89 65 70 31 60 40 60 81 19 68 64 36 74 45 69 36 38 25 39
5	50 44 66 44 21 22 66 22 15 86 96 24 40 14 51 31 73 91 61 19 78 60 73 99 84	66 06 58 05 62 26 63 75 41 99 23 22 30 88 57 60 20 72 93 48 43 89 94 36 45	68 15 54 35 02 58 42 36 72 24 95 67 47 29 83 98 57 07 23 69 56 69 47 07 41	42 35 48 96 32 58 37 52 18 51 94 69 40 06 07 65 95 39 69 58 90 22 91 07 12
6	84 37 90 61 56 36 67 10 08 23 07 28 59 07 48 10 15 83 87 60 55 19 68 97 65	70 10 23 98 05 98 93 35 08 86 89 64 58 89 75 79 24 31 66 56 03 73 52 16 56	85 11 34 76 60 99 29 76 29 81 83 85 62 27 89 21 48 24 06 93 00 53 55 90 27	76 48 45 34 60 33 34 91 58 93 30 14 78 56 27 91 98 94 05 49 33 42 29 38 87
7	53 81 29 13 39 51 86 32 68 92 35 91 70 29 13 37 71 67 95 13 93 66 13 83 27	35 01 20 71 34 33 98 74 66 99 80 03 54 07 27 20 02 44 95 94 92 79 64 64 72	62 33 74 82 14 40 14 71 94 58 96 94 78 32 66 64 85 04 05 72 28 54 96 53 84	53 73 19 09 03 45 94 19 38 81 50 95 52 74 33 01 32 90 76 14 48 14 52 98 94
8	02 96 08 45 65 49 83 43 48 35 84 60 71 62 46 18 17 30 88 71 79 69 10 61 78	13 05 00 41 84 82 88 33 69 96 40 80 81 30 37 44 91 14 88 47 71 32 76 95 62	93 07 54 72 59 72 36 04 19 76 34 39 23 05 38 89 23 30 63 15 87 00 22 58 40	21 45 57 09 77 47 45 15 18 60 25 15 35 71 30 56 34 20 47 89 92 54 01 75 25

Per usare bene la tabella occorre selezionare, per sorteggio, il blocco riga/colonna da cui iniziare e poi prestabilire come continuare

Numeri pseudo-casuali

I numeri pseudo-casuali sembrano prodotti dalla sorte, ma sono noti a priori

$$X_k = Resto\left[aX_0 + \frac{(a^k - 1)c}{(a-1)}, m\right] \qquad [\ ]\text{=parte intera}$$

Basta conoscere il primo e tutti gli altri sono noti.

La sequenza è ciclica: dopo "m" valori i numero si ripetono nello stesso ordine

Il periodo "m" deve essere grande rispetto al campione da estrarre: Regola di Ripley

$$m \geq 200n^2 \Rightarrow se \quad n = 10'000 \quad m \geq 20'000'000'000 > 2^{31}$$

Esempio

$$X_{i+1} \equiv (aX_i + c)\text{Mod } m \qquad \text{con} \begin{cases} a = 293 \\ c = 1 \\ m = 1024 \\ X_1 = 68 \end{cases}$$

68	469	202	819	352	737	902	95	188	813	642	715	600	697	446
631	564	389	314	867	80	913	246	399	172	221	242	251	840	361
302	423	36	309	426	915	832	65	614	703	156	653	866	811	56
25	158	215	532	229	538	963	560	241	982	1007	140	61	466	347
296	713	14	7	4	149	650	1011	288	417	326	287	124	493	66
907	536	377	894	823	500	69	762	35	16	593	694	591	108	925

Se è necesario un numero pseudo casuale compreso tra zero ed uno basterà dividere il risultato dello schema per il modulo "m"

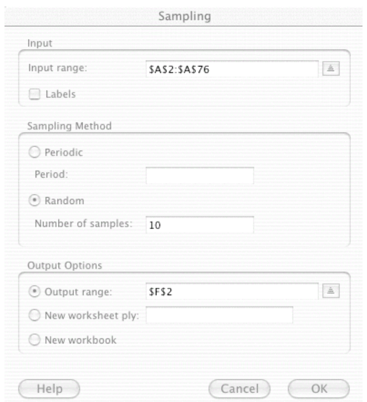
Codice	Lista
Urto	77.08
Yama	78.56
Xilo	79.55
Wide	80.87
Zone	81.43
Dior	82.47
Idea	82.78
Urna	83.65
Mola	84.43
Ecos	85.31
Dato	86.42
Ceca	86.78
Alea	87.99
Tino	88.23
Erto	89.47
Mega	90.61
Doge	92.22
Alfa	92.23
Page	92.32
Foce	92.49
Jana	93.28
Teco	93.50
Oasi	94.12
Daga	94.25
Diga	94.75
Remo	95.28
Host	95.36
Sale	95.45
Biro	96.12
West	96.21
Riso	96.31
Cala	96.64
Indy	96.84
Posa	97.28
Demo	97.48
Duna	97.50
Quod	97.71
Wild	98.05
Yack	98.11
Nora	98.18
Vice	99.37
Nord	99.74
Seta	99.97
Home	100.10
Remi	100.69
Luce	101.24
Unno	101.52
Dote	101.69
Sito	101.88
Zeta	101.89
Vela	102.09
Beta	102.25
Paga	102.35
Tura	103.02
Nave	103.50
Xeno	103.85
Gelo	103.97
Arca	104.97
Raso	105.52
Zara	106.86
Sula	106.92
Orso	107.24
Jena	108.17
Mina	109.27
Muta	110.13
Nero	111.65
Rada	112.81
Arpa	112.98
Lena	113.98
Giro	114.74
Hera	115.91
Jole	116.53
Oboe	116.96
Polo	118.61
Xmas	119.44

Excel: C.C.S. con Reimmissione

EXCEL ha una procedura per il campionamento con reimmissione

<u>Strumenti: Analisi dei dati: Campionamento</u>

Dopo aver indicato l'intervallo di input e il numero di unità da campionare Excel costruisce un campione casuale con reimmissione nella zona di output richiesta



Unità campionate	
Mola	84.43
Ceca	86.78
Jana	93.28
Remo	95.28
Remo	95.28
Remo	95.28
Demo	97.48
Home	100.1
Tura	103.02
Jole	116.53

Codice	Lista
Urto	77.08
Yama	78.56
Xilo	79.55
Wide	80.87
Zone	81.43
Dior	82.47
Idea	82.78
Urna	83.65
Mola	84.43
Ecos	85.31
Dato	86.42
Ceca	86.78
Alea	87.99
Tino	88.23
Erto	89.47
Mega	90.61
Doge	92.22
Alfa	92.23
Page	92.32
Foce	92.49
Jana	93.28
Teco	93.50
Oasi	94.12
Daga	94.25
Diga	94.75
Remo	95.28
Host	95.36
Sale	95.45
Biro	96.12
West	96.21
Riso	96.31
Cala	96.64
Indy	96.84
Posa	97.28
Demo	97.48
Duna	97.50
Quod	97.71
Wild	98.05
Yack	98.11
Nora	98.18
Vice	99.37
Nord	99.74
Seta	99.97
Home	100.10
Remi	100.69
Luce	101.24
Unno	101.52
Dote	101.69
Sito	101.88
Zeta	101.89
Vela	102.09
Beta	102.25
Paga	102.35
Tura	103.02
Nave	103.50
Xeno	103.85
Gelo	103.97
Arca	104.97
Raso	105.52
Zara	106.86
Sula	106.92
Orso	107.24
Jena	108.17
Mina	109.27
Muta	110.13
Nero	111.65
Rada	112.81
Arpa	112.98
Lena	113.98
Giro	114.74
Hera	115.91
Jole	116.53
Oboe	116.96
Polo	118.61
Xmas	119.44

C.C.S. Senza Reimmissione

Si può procedere nel modo seguente

1)Si pone a fianco delle righe "popolazione" una nuova colonna SELETTORE in cui sia inserita la funzione CASUALE().

2) Si riordinano le righe in base al SELETTORE (ascendente o discendente non importa)

3) Le unità che si trovano nelle prime "n" posizioni sono il CCS-SR.

Codice	Lista	Selettore
Host	95.36	0.339044806
Demo	97.48	0.91048498
Beta	102.25	0.477187422
Orso	107.24	0.685980471
Page	92.32	0.256799318
Indy	96.84	0.309094032
Yack	98.11	0.239753839
Unno	101.52	0.469452099
Sula	106.92	0.554879814
Sito	101.88	0.721677718

E' poco probabile che due di questi valori siano uguali(se n<32768)