• Distribuzione ipergeometrica

$\#$ pesci in un lago = ?

Faccio un esperimento

- pesco 1000 pesci, gli metto un anello e li rimetto nel ~~lago~~ lago
- ... ripesco altri 1000 pesci e osservo che 100 di questi hanno l'anello

Se volessi fare affermazioni certe

$$\# \text{ pesci} \geq 1900$$

non è tanto interessante.

Qual'è il numero di pesci più verosimile ? (alla luce dell'esperimento fatto)

Senza sapere nulla

$N = \#$ pesci (incognita)

dopo la prima pescata

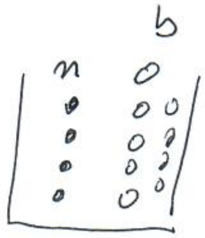
$$\underset{\#\text{ pesci}}{N} : \underset{\#\text{ pesci con anello}}{1000} = \underset{\#\text{ pesci con anello}}{1000} : \underset{\#\text{ pesci con anello presi il secondo giro}}{100}$$

$$\Rightarrow N = 1000 \cdot 10 = 10^4$$

Vedremo che questo è effettivamente il numero di pesci più "verosimile"

Formulazione come problema di estrazione da urne

Il primo giro serve solo a preparare
Al momento del secondo



$N = n + b$ = # pesci nel lago

b = # pesci con anello
n = # " senza anello

estraggo k palline (non ordinate, senza rimpiazzo)

$\mathbb{P}$(h delle k palline estratte sono nere) = ? $\quad h = 0, \dots, \min\{k, n\}$

OSS se fosse con rimpiazzo sarebbe come schema di Bernoulli

costruisco uno schema a prob. unif.
numero le palline $\underbrace{1, 2, \dots, n}_{\text{nere}}, \underbrace{n+1, \dots, n+b}_{\text{bianche}}$ ($n+b = N$)

Ω = campionamento non ordinato senza rimpiazzo $\quad |\Omega| = \binom{n+b}{k}$

prob. uniforme

che evento mi interessa

$A = \{$ delle k estratte h sono nere $\}$

Facciamo prima uno più facile

$$B = \left\{ \begin{array}{l} \text{estraggo } 1, 2, \dots, h \\ \text{ma non } h+1, h+2, \dots, n \end{array} \right\}$$

$$|B| = \binom{b}{k-h}$$

ora, come unione di eventi disgiunti,

$$|A| = \binom{n}{h} |B|$$

quindi

$$\mathbb{P}(A) = \frac{\binom{n}{h}\binom{b}{k-h}}{\binom{n+b}{k}}$$

convengo che $= 0$ se $h > n$ oppure $h > k$

- Ritorniamo al problema iniziale

$n = 1000$ = # pesci con anello (noto)

$k = 1000$ = # " presi la seconda volta (noto)

$n + b$ (o equiv. b) ignoto

$h = 100$ evento osservato

Come decido $n + b = N$ "verosimile" ?

criterio massima verosimiglianza

Scelgo il valore del parametro incognito ($N = n+S$) che max la prob. dell'evento osservato ($h = 100$)

ora è un problema di matematica preciso

n, k, h fissati $\quad n+S = N$ incognito

guarda la funzione

$$P_{n,k,h}(N) = \frac{\binom{n}{h}\binom{N-n}{k-h}}{\binom{N}{k}}$$

dove assume max ?

se crediamo alla proporzione iniziale dovrà capitare per

$$N = n\,\frac{k}{h}$$

(parti intere da aggiustare)

come si dimostra ?

derivata ? Ma N è solo intero

in effetti è l'idea giusta

CAPIAMO SE

$$p(N) \gtrless p(N-1) \quad \text{quando capita}$$

conviene fare il rapporto invece:

$$\cancel{\frac{p(N)}{p(N-1)} > 1}$$

per quali N accade $\dfrac{p(N)}{p(N-1)} > 1$?

(questo ci dice quando è crescente e quando decrescente)

$$\frac{p(N)}{p(N-1)} = \frac{\binom{n}{h}\binom{N-n}{k-h}}{\binom{N}{k}} \; \frac{\binom{N-1}{k}}{\binom{n}{h}\binom{N-1-n}{k-h}}$$

$$= \frac{\dfrac{(N-n)!}{(k-h)!\,(N-n-k+h)!}}{\dfrac{N!}{k!\,(N-k)!}} \cdot \frac{\dfrac{(N-1)!}{k!\,(N-1-k)!}}{\dfrac{(N-1-n)!}{(k-h)!\,(N-1-n-k+h)!}}$$

$$= \frac{(N-n)/(N-n-k+h)}{N/(N-k)} = \frac{N-n}{N}\;\frac{N-k}{N-n-k+h}$$

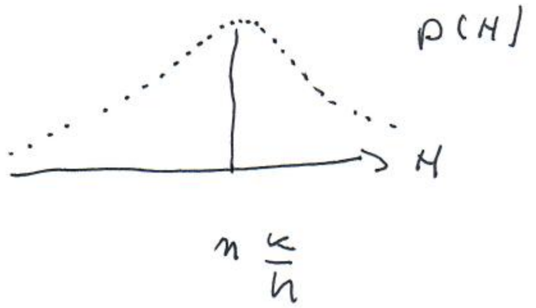
$$\frac{p(N)}{p(N-1)} > 1 \iff \frac{N-n}{N}\;\frac{N-k}{N-n-k+h} > 1$$

$$\iff (N-n)(N-k) > N(N-n-k+h)$$

$$\iff N^2 - N(n+k) + nk > N^2 - N(n+k) + Nh$$

$$\iff nk > Nh \iff N < n\frac{k}{h}$$

ci ha detto bene



[parti intere da sistemare]

Abbiamo risolto un problema di statistica matematica

" [In questo contesto] lo stimatore di max verosimiglianza per N è

$n\frac{k}{h}$ "

• Mendace assiomi probabilità

In situazioni un po' più raffinite di Ω finito avremo necessità (soprattutto volonté) di fare operazioni infinite sugli eventi. $\mathcal{A}$ algebra e $\mathbb{P}$ additive non ci basterà più.

Esempio Notte di capodanno

(A) ½ notte − 1 min metto in urna palline 1, ..., 10 ed estraggo la 10

½ notte − ½ min " " " 11, ..., 20 ed estraggo la 20

⋮

½ notte − ¼ min ... 21, ..., 30

... 30

palline nell'urna a ½ notte in punto $= \infty$

(ovvero più di ogni intero fissato)

(B) ½ notte − 1 min metto in un urna 1, ..., 10 ed estraggo 1

½ notte − $\frac{1}{2}$ min " " " 11, ..., 20

⋮ ... 2

palline nell'urna a ½ notte in punto $= 0$

1 non c'è, 2 non c'è, ...

[bisogna adattarsi ai fatti della vita]

(C) $1/2$ notte $- 1$ min metto palline $1, \ldots, 10$
e ne estraggo una a caso

$\frac{1}{2}$ notte $- \frac{1}{2}$ min " " $11, \ldots, 20$
e ne estraggo una a caso

$\vdots$

\# palline nell'urna
a $\frac{1}{2}$ notte in punto $=$?

Domanda chiave

$E = \{$ pallina 1 è rimasta
nell'urna a $1/2$ notte in punto $\}$

$$\mathbb{P}(E) \begin{matrix} = 0 \\ > 0 \end{matrix} \quad ?$$

$E_n = \{$ pallina 1 è rimasta nell'urna
dopo le prime n estrazioni $\}$

oss
$$E = \bigcap_{n=1}^{\infty} E_n \qquad E_{n+1} \subset E_n$$

$$\mathbb{P}(E_1) = \frac{9}{10}$$

$$\mathbb{P}(E_2) = \frac{9}{10} \cdot \frac{18}{19}$$

infatti $\Omega^{(2)} = \{ (\omega_1, \omega_2)$ prob. unif
$\omega_1 = 1, \ldots, 10$
$\omega_2 = 1, \ldots, 20 \quad \omega_2 \neq \omega_1 \}$

$|\Omega^{(2)}| = 10 \cdot 19 \qquad |E_2| = 9 \cdot 18$

$$\mathbb{P}(E_n) = \frac{9}{10} \frac{18}{19} \cdots \frac{9n}{9n+1}$$

Poiché $E_{n+1} \subset E_n$ $\quad \mathbb{P}(E_n) \searrow$

$$\lim_{n \to \infty} \mathbb{P}(E_n) = \prod_{i=1}^{\infty} \frac{9i}{9i+1} = ?$$

Ex

$$\prod_{i=1}^{\infty} \frac{9i}{9i+1} = 0$$

dim

è equiv. a

$$\lg \prod_{i=1}^{\infty} \frac{9i}{9i+1} = -\infty \iff \sum_{i=1}^{\infty} \lg \frac{9i}{9i+1} = -\infty$$

ora

$$\lg \frac{9i}{9i+1} = \lg\left(1 - \frac{1}{9i+1}\right)$$

$$\left| \lg\left(1 - \frac{1}{9i+1}\right) + \frac{1}{9i+1} \right| \leq \frac{C}{(9i+1)^2} \quad i \geq 1$$

[via Lagrange]

$$\sum_i \frac{1}{9i+1} = +\infty$$

$$\sum_i \frac{1}{(9i+1)^2} < +\infty$$

$$\Rightarrow \sum_{i=1}^{\infty} \lg\left(1 - \frac{1}{9i+1}\right) = -\infty$$

Vogliamo concludere

$$\mathbb{P}(E) = \lim_{n \to \infty} \mathbb{P}(E_n) = 0$$

D) Aggiungete palline in modo geometrico non aritmetico

$\frac{1}{2}$ volte - 1 min   mettete 1, 2 e estraete una a caso

$\frac{1}{2}$ volte - $\frac{1}{2}$ min   " 3, 4, 5, 6

$\frac{1}{2}$ volte - $\frac{1}{4}$ min   " " " " " 7 - 14 e togliete una a caso

⋮

ex $E_n = \{$ la pallina 1 è ancora nell'urna dopo n estrazioni $\}$

dim che

$$\lim_{n\to\infty} \mathbb{P}(E_n) > 0$$

Assiomi emendati

SPAZIO DI PROB. $(\Omega, \mathcal{F}, \mathbb{P})$

DOVE

- Ω INSIEME (collezione degli eventi elementari)

- $\mathcal{F}$ ~~ALGEBRA~~ EVENTI
  σ-ALGEBRA

dove σ-algebra è una famiglia di sottoinsiemi di Ω chiusa rispetto a

- passaggio al complementare
- unioni numerabili
- intersezioni numerabili

ovvero

$$A_i \in \mathcal{F} \quad i = 1, \dots, n, \dots$$

$$\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}, \quad \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

[poiché è chiusa rispetto a $\dots^c$, basta unioni num. o intersezioni num.]

(o) $\mathbb{P} : \mathcal{F} \longmapsto [0,1]$

~~ADDITIVA~~

σ-ADDITIVA

ovvero

se $A_i \in \mathcal{F}$ $\quad i = 1, \dots, n, \dots$

disgiunti $\quad (i \neq j \Rightarrow A_i \cap A_j = \phi)$

$$\mathbb{P}\Big(\underbrace{\bigcup_{i=1}^{\infty} A_i}\Big) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

$\rightsquigarrow \in \mathcal{F}$ ~~PER LE PROPRIETÀ DI $\mathcal{F}$~~
poiché $\mathcal{F}$ è σ-ALG.

— o —

Risposta all'esempio iniziale

Invero (prossima volta)

$$E_{n+1} \subset E_n$$

$$E = \bigcap_{n=1}^{\infty} E_n$$

$$\mathbb{P} \ \sigma\text{-ADD} \Rightarrow \mathbb{P}(E) = \lim_n \mathbb{P}(E_n)$$