FUNZIONI DI PROBABILITÀ DI VARIABILI DISCRETE

CdS Fisica

Laboratorio Meccanica e Termodinamica

Funzioni di Probabilità di Variabili Discrete

- Sono frequenti situazioni nelle quali si devono misurare **quantità come flussi e densità di "oggetti"** (particelle o altro).

  Ad esempio:
  - flusso di raggi cosmici: numero di raggi cosmici per unità di area e di tempo;
  - densità di micro-particelle in sospensione fluida (particolato atmosferico);
  - attività di una sorgente radioattiva: numero medio di decadimenti per unità di tempo.

- In questi casi, la grandezza da misurare, o contare, è un **numero intero**, che può variare da misura a misura. Si tratta quindi di una **variabile casuale discreta**.
- Effettuando un'unica misura, quale **incertezza si deve associare al conteggio ottenuto ?**

Funzioni di Probabilità di Variabili Discrete (II)

- Le variabili casuali discrete (cioè non continue) hanno una trattazione statistica simile a quella delle variabili casuali continue, che abbiamo considerato fino ad ora. Le loro funzioni (o distribuzioni) di probabilità **non sono però delle densità**, dato il carattere discreto della variabile;

- Le distribuzioni di probabilità di variabile discreta di maggiore interesse per la fisica sono la **distribuzione binomiale** e la **distribuzione di Poisson**;

- La distribuzione più semplice è però quella **uniforme** (lancio di una moneta o di un dado non truccati, numero di bambini nati in un reparto maternità…).

Funzioni di Probabilità di Variabili Discrete (III)

Stesse proprietà delle distribuzioni continue:

- Normalizzazione $$\boxed{\sum_{k=0}^{N} P(k) = 1}$$

- Media $$\boxed{\bar{k} = \sum_{k=0}^{N} kP(k)}$$

- Varianza $$\boxed{\sigma_k^2 = \sum_{k=0}^{N} (k - \bar{k})^2 P(k) = \overline{k^2} - \bar{k}^2}$$

Distribuzione Uniforme

- Variabile discreta k nel lancio di un dado;

- $P(k) = \frac{1}{6}$ con $k = 1,2, \dots 6$

- $P(k)$ è una probabilità con una distribuzione uniforme; non è una densità di probabilità come invece avviene per le variabili continue;

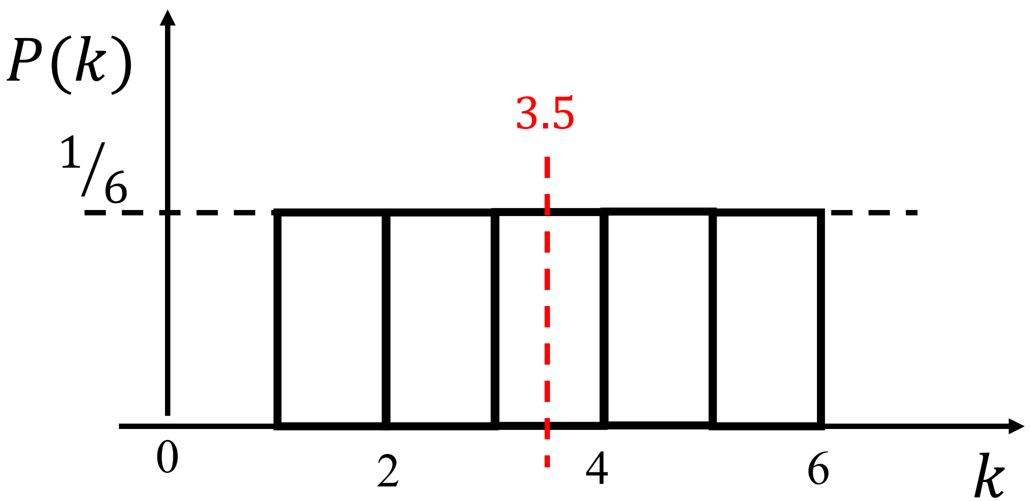
- Condizione di normalizzazione:
  $\sum_{i=1}^{6} P(k) = 6 \times \frac{1}{6} = 1$
  corrisponde all'integrale nel caso di variabile continua;

Distribuzione Uniforme (II)

Valor medio: $\bar{k}$ atteso

$$\bar{k} = \sum_{k=1}^{6} kP(k) = \sum_{k=1}^{6} k\frac{1}{6} = \frac{1}{6}\boxed{\sum_{k=1}^{6} k} = \frac{1}{6}\frac{6\cdot 7}{2} = \frac{7}{2} = 3.5$$

Formula di Gauss $\boldsymbol{s = 1 + 2 + \cdots n = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}}$



Distribuzione Uniforme (III)

Varianza:

$$\sigma_k^2 = \sum_{k=1}^{6} (k - \bar{k})^2 P(k) = \overline{k^2} - \bar{k}^2 \qquad \bar{k} = 3.5$$

$$\overline{k^2} = \sum_{k=1}^{6} k^2 P(k) = \sum_{k=1}^{6} k^2 \frac{1}{6} = \frac{1}{6} \boxed{\sum_{k=1}^{6} k^2} = \frac{1}{6} \frac{6 \cdot 7 \cdot 13}{6} = 15.2$$

$$\boldsymbol{s = 1^2 + 2^2 + \cdots n^2 = \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}}$$

$$\sigma_k = \sqrt{\overline{k^2} - \bar{k}^2} = \sqrt{15.2 - 3.5^2} = 1.7$$

Calcolo Combinatorio: Disposizioni

Il numero di **modi diversi di disporre k oggetti, presi da un insieme di n oggetti**, è chiamato numero di **disposizioni** ed è indicato con $D_{n,k}$ e risulta che:

$$\boxed{D_{n,k} = \frac{n!}{(n-k)!}}$$

Esempio

Quante sono le possibili configurazioni del podio (oro, argento, bronzo) in una gara di 100 m a otto finalisti ?

(risposta: 336)

Disposizioni (II)

Prendiamo il caso di 3 oggetti (a,b,c) che voglio disporre a gruppi di due: $n = 3$ e $k = 2$

Esistono 6 disposizioni possibili di 2 oggetti presi da un gruppo di 3: (a,b), (b,a), (a,c), (c,a), (b,c), (c,b).

$$D_{n,k} = \frac{n!}{(n-k)!} \Longrightarrow D_{3,2} = \frac{3!}{(3-2)!} = \frac{3!}{1!} = 6$$

Una volta scelto il primo oggetto, la scelta del secondo oggetto può avvenire solo fra i restanti n-1.

Disposizioni (III)

Generalizzando:

$$\underbrace{n}_{\text{Scelta del primo}}\underbrace{(n-1)}_{\text{Scelta del secondo fra i (n-1) restanti}}\underbrace{(n-2)}_{\text{Scelta del terzo fra i (n-2) restanti}}\dots\underbrace{(n-k+1)}_{\text{Scelta del penultimo}} =$$

Ultimo è ovvio

$$n(n-1)(n-2)\dots(n-k+1)\frac{(n-k)(n-k-1)\dots 1}{(n-k)(n-k-1)\dots 1} =$$

$$= \frac{n!}{(n-k)!}$$

Cvd

Calcolo Combinatorio: Permutazioni

Le **permutazioni** P_n di n oggetti sono costituite dai **modi diversi di ordinare tali oggetti**. Risulta quindi che $P_n = D_{n,n}$, cioè:

$$\boxed{P_n = n!}$$

Esempio

In quanti possibili modi diversi gli otto finalisti possono occupare le otto corsie ?

(risposta: 40320)

Calcolo Combinatorio: Combinazioni Senza Ripetizioni

Il numero di **modi diversi di disporre k oggetti, presi da un insieme di n oggetti, senza distinguere i gruppi** di k oggetti che si differenziano solo per l'ordine nel quale gli oggetti si presentano, è chiamato numero di **combinazioni** (**disposizioni senza ripetizione**) ed è indicato con $C_{n,k}$.

$$\boxed{C_{n,k} = \frac{n!}{(n-k)!\,k!} = \binom{n}{k}}$$

Coefficiente binomiale

Esempio

In una classe di 12 studenti si formano 4 gruppi di 3 studenti per sessioni di laboratorio. Quante configurazioni sono possibili per il primo dei 4 gruppi ? (risposta: 220)

Calcolo Combinatorio: Combinazioni Con Ripetizioni

Quando l'ordine non è importante, ma è possibile avere componenti ripetute, si parla di **combinazioni con ripetizione**: il numero di **modi diversi di disporre k oggetti, presi da un insieme di n oggetti, dove uno stesso oggetto può essere ripetuto fino a k volte**

$$C_{n,k} = \binom{n+k-1}{k}$$

Distribuzione Binomiale

- Consideriamo un **processo casuale con due soli possibili esiti.**
- Se indichiamo con **p la probabilità di una delle due uscite ("successo"), allora la probabilità dell'altra uscita ("insuccesso") sarà q = 1 – p.**
- Ad esempio prendiamo il lancio di un dado, essendo "successo" l'uscita del numero "5". In questo caso, $p = 1/6$ e $q = 5/6$.

Distribuzione Binomiale (II)

- Consideriamo adesso l'esperimento di lanciare n = 10 volte il dado, calcolando la probabilità di ottenere k = 2 volte il "successo".

- L'esperimento può essere visualizzato con una sequenza di dieci numeri; la probabilità che 2 di questi siano dei "5" può essere determinata in base a quanto conosciamo sul calcolo delle probabilità e sul calcolo combinatorio:

- $P = \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8 + \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8 + \cdots \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8$

  Somma di $C_{10,2} = \frac{n!}{(n-k)!k!} = \frac{10!}{(10-2)!2!} = 45$ termini uguali.

  Quindi $P = 45 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8 = 0.29$

Distribuzione Binomiale (II) bis

$5, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$5, \bar{5}, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$5, \bar{5}, \bar{5}, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$5, \bar{5}, \bar{5}, \bar{5}, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, 5, \bar{5}, \bar{5}, \bar{5}$

$5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, 5, \bar{5}, \bar{5}$

$5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, 5, \bar{5}$

$5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, 5$

$\bar{5}, 5, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$\bar{5}, 5, \bar{5}, 5, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}, \bar{5}$

$\ldots$

Distribuzione Binomiale (III)

In generale, indicando con p la probabilità del "successo" e ripetendo n volte il processo casuale, la probabilità di ottenere k successi (con $0 \leq k \leq n$) sarà data dalla relazione:

$$\boxed{P(k;n,p) = \binom{n}{k} p^k q^{n-k}}$$

Distribuzione Binomiale

Infatti l'evento = "k successi su n prove" è l'unione di $C_{n,k}$ eventi mutuamente escludentisi.
A sua volta, ognuno dei $C_{n,k}$ termini è l'intersezione di n eventi indipendenti, k di esito positivo e (n - k) di esito negativo.
La distribuzione può essere rappresentata da un grafico, con k (la variabile casuale) in ascissa e P in ordinata, essendo n e p due parametri.

Distribuzione Binomiale (IV)

La distribuzione binomiale soddisfa le seguenti proprietà:

- la distribuzione è **normalizzata** a 1 (come deve, essendo una distribuzione di probabilità)

- la media della variabile k è $\bar{k} = np$
  la varianza è $\sigma^2 = npq = np(1-p)$

- la distribuzione è **simmetrica solo se p = 0.5**

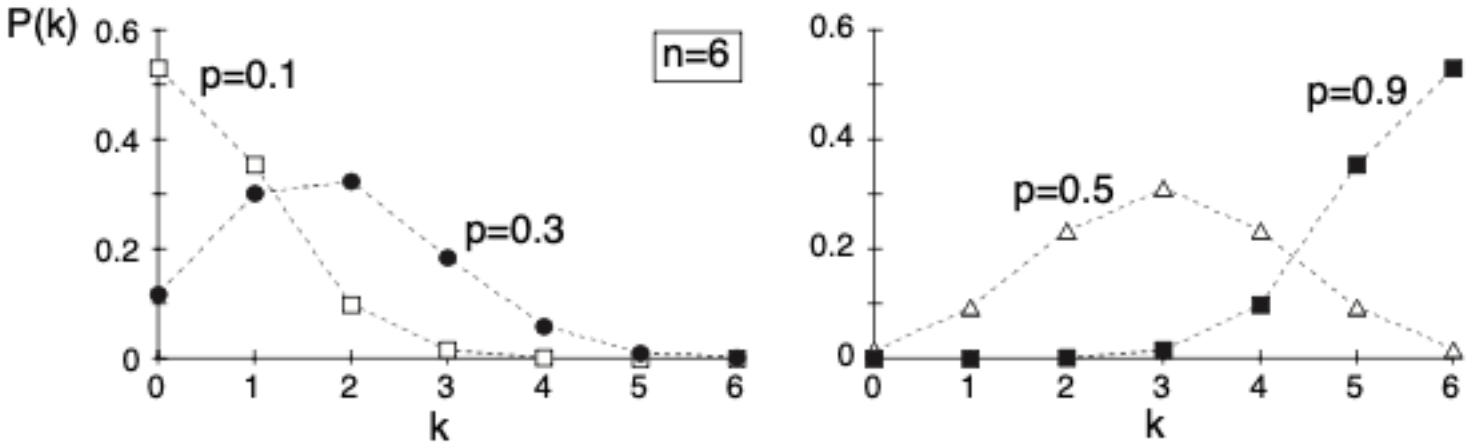
- la distribuzione **tende ad una gaussiana** per grandi valori di n

$$\boxed{(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}}$$

Formula del Binomio

Distribuzione Binomiale (V)

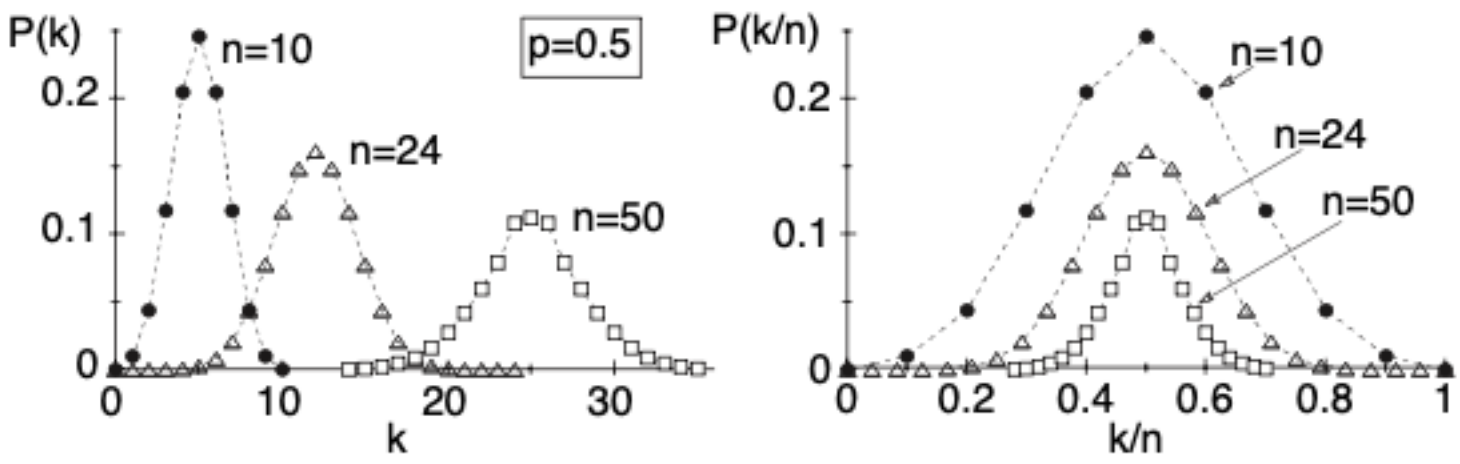
La distribuzione Binomiale è **simmetrica solo se p = 0.5**



Distribuzione binomiale per n=6 e diversi valori di *p*.
Per $p = 0.5$, la distribuzione è simmetrica in $k = np = 6*0.5 = 3$

Distribuzione Binomiale (VI)

La distribuzione Binomiale **tende ad una gaussiana** per grandi valori di n.



Distribuzione binomiale per $p = 0.5$ e diversi valori di n.

Distribuzione Binomiale: Normalizzazione

La distribuzione binomiale è normalizzata a 1.

Dimostrazione

$$P(k; n, p) = \binom{n}{k} p^k q^{n-k}$$

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = [\cancel{p} + (1-\cancel{p})]^n = 1^n = 1$$

Cvd

Formula del binomio

In una densità di probabilità di variabile continua avrei avuto

$$\int_{-\infty}^{+\infty} \phi(x)dx$$

Distribuzione Binomiale: Media

$\bar{k} = np$

Dimostrazione

$$P(k; n, p) = \binom{n}{k} p^k q^{n-k}$$

$$\bar{k} = \sum_{k=0}^{n} kP(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} \cancel{k} \frac{n!}{(n-k)!\, \underbrace{\cancel{k}!}_{=(k-1)!}} p^k q^{n-k}$$

Per k=0, il termine della sommatoria è nullo, possiamo quindi partire con k=1

$$= \sum_{k=1}^{n} \boxed{np} \frac{(n-1)!}{(n-k)!\,(k-1)!} p^{k-1} q^{n-k} = np \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!\,(k-1)!} p^{k-1} q^{n-k}$$

Distribuzione Binomiale: Media (II)

$$\bar{k} = np \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!\,(k-1)!} p^{k-1} q^{n-k}$$

Cambio di variabili: $h = k - 1,\ m = n - 1$

$\Longrightarrow k = h + 1,\ n = m + 1,\ n - k = m + 1 - (h + 1) = m - h$

Se $k = 1 \Longrightarrow h = 0$; Se $k = n \Longrightarrow h = n - 1 = m$

$$\bar{k} = np \boxed{\sum_{h=0}^{m} \frac{m!}{(m-h)!\,h!} p^h q^{m-h}} = np(p+q)^m = np1 = np$$ Cvd

= 1 normalizzazione
(formula del binomio)

Distribuzione Binomiale: Varianza

$\sigma^2 = npq = np(1-p)$

Dimostrazione

$$P(k;n,p) = \binom{n}{k} p^k q^{n-k}$$

$\sigma^2 = \overline{k^2} - \bar{k}^2$ Dove $\bar{k} = np$ da dimostrazione precedente

$$\overline{k^2} = \sum_{k=0}^{n} k^2 P(k) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} k^{\cancel{2}} \frac{n!}{(n-k)!\,\underbrace{\cancel{k}!}_{=(k-1)!}} p^k q^{n-k}$$

$$= \sum_{k=1}^{n} \boxed{np} k \frac{(n-1)!}{(n-k)!\,(k-1)!} p^{k-1} q^{n-k} = np \sum_{k=1}^{n} k \frac{(n-1)!}{(n-k)!\,(k-1)!} p^{k-1} q^{n-k}$$

Distribuzione Binomiale: Varianza (II)

$$\overline{k^2} = np\sum_{k=1}^{n} k\frac{(n-1)!}{(n-k)!\,(k-1)!}p^{k-1}q^{n-k}$$

Cambio di variabili: $h = k - 1$, $m = n - 1$

$\Longrightarrow k = h + 1, n = m + 1, n - k = m + 1 - (h + 1) = m - h$

Se $k = 1 \Longrightarrow h = 0$; Se $k = n \Longrightarrow h = n - 1 = m$

$$\overline{k^2} = np\sum_{h=0}^{m}(h+1)\frac{m!}{(m-h)!\,h!}p^h q^{m-h} =$$

Sommatoria di una somma = somma di sommatorie

$$= np\left(\boxed{\sum_{h=0}^{m} h\frac{m!}{(m-h)!\,h!}p^h q^{m-h}} + \boxed{\sum_{h=0}^{m}\frac{m!}{(m-h)!\,h!}p^h q^{m-h}}\right) = np(mp+1)$$

Media di una binomiale con variabile casuale h $\Longrightarrow$ mp

Normalizzazione $\Longrightarrow (p+q)^m$=1

Distribuzione Binomiale: Varianza (III)

Cambio di variabili: $m = n - 1$

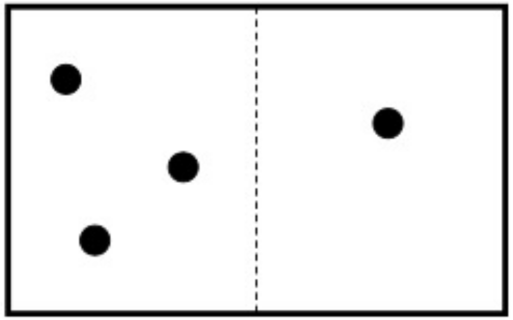
$$\overline{k^2} = np(mp+1) = np[(n-1)p+1] = np(np-p+1)$$
$$= (np)^2 - np^2 + np$$

$$\sigma^2 = \overline{k^2} - \bar{k}^2 = [(np)^2 - np^2 + np] - (np)^2 = -np^2 + np$$

$$= np(1-p) = npq$$ Cvd

Esercizio 1

Quattro molecole si trovano in un contenitore idealmente diviso in due parti uguali (chiamate "sinistra" e "destra"). Qual è la probabilità che, ad un certo istante, k molecole ($k = 0, 1, 2, 3, 4$) si trovino nella parte "sinistra" ?



Esercizio 2

Un apparato è alimentato tramite generatori di corrente. Per ridondanza, ci sono 3 generatori, tra loro indipendenti ed identici. Per funzionare, l'apparato deve essere alimentato da almeno due dei tre generatori.

Se ognuno degli alimentatori eroga corrente in modo corretto nell'80% del tempo, qual è la probabilità che l'apparato funzioni, ad un dato istante ?

Esercizio 3

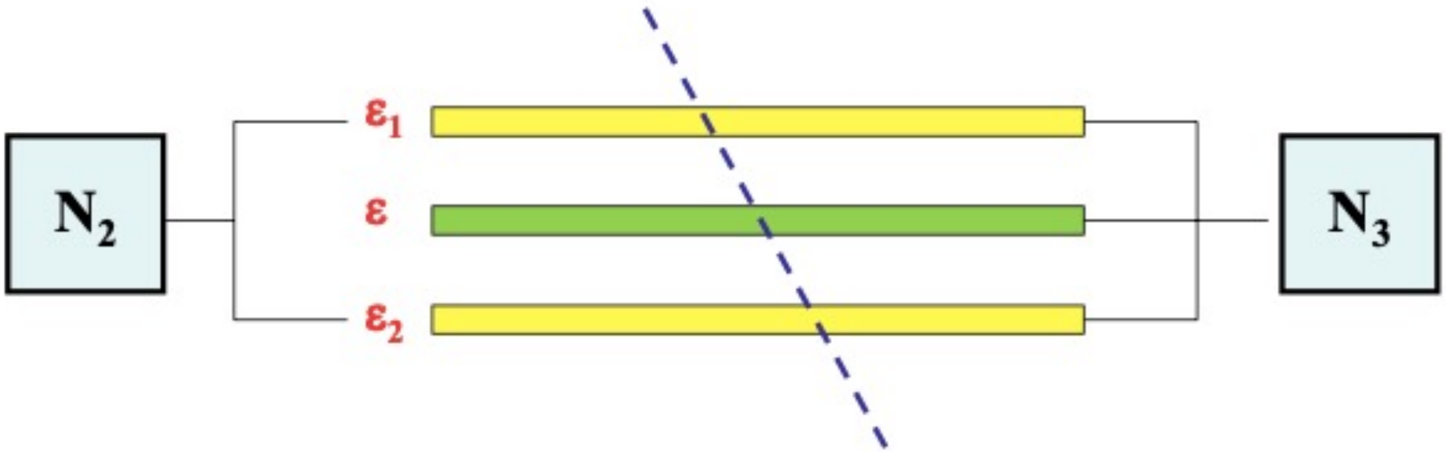
In una industria, si produce un certo tipo di elettrodomestico; il 4% dei pezzi risulta essere difettoso.

1) Se vengono selezionati, a caso, 8 pezzi, qual è la probabilità che nessuno sia difettoso ?

2) Il valore del 4% era stato ottenuto analizzando 200 pezzi e trovandone 8 difettosi. Qual è l'incertezza (pari ad una deviazione standard) da attribuire a tale probabilità ?

Esercizio 4

Si vuole determinare l'efficienza ε di un contatore di raggi cosmici, e la sua incertezza $\Delta\varepsilon$. Allo scopo, si realizza la configurazione mostrata in figura, utilizzando due contatori "ausiliari" di efficienze (incognite) ε_1 ed ε_2.
In un certo intervallo di tempo, si misurano $N_2 = 150$ coincidenze (= evento intersezione) tra i due contatori ausiliari ed $N_3 = 135$ coincidenze tra i tre contatori. Quanto valgono ε e l'incertezza $\Delta\varepsilon = \sigma_\varepsilon$?



Esercizio 5

Per contare un certo tipo di particelle, emesse da una sorgente, si usa un *detector* composto da due diversi rivelatori, funzionanti "a cascata" e in modo indipendente.

In un determinato intervallo di tempo, il rivelatore A conta $n_A = 1849$ volte mentre il rivelatore B conta $n_B = 2007$ volte.

In coincidenza, i conteggi sono $n_{AB} = 1237$.

(i) Qual è il numero N di particelle emesse dalla sorgente in quell'intervallo di tempo ?

(ii) Se definiamo ε l'efficienza totale del *detector* (= probabilità che almeno uno dei due rivelatori si attivi al passaggio di una particella), quanto vale ε ?

(iii) Quanto vale l'incertezza su ε, cioè σ_ε ?

Efficienza di Conteggio

In generale, definiamo **l'efficienza di un conteggio**:

Sia *N* il numero di particelle che attraversano un rivelatore.
Sia *n* in numero di particelle effettivamente rivelate.

Si definisce efficienza del rivelatore $\boldsymbol{\varepsilon = \frac{n}{N}}$

L'efficienza ha una certa incertezza, che è calcolata su un certo numero di misure.

L'incertezza sull'efficienza è un errore casuale: $\boldsymbol{\Delta\varepsilon = \sigma_{\varepsilon}}$

Efficienza di Conteggio (II)

$$\varepsilon = \frac{n}{N}$$

Immagino di tenere N costante in diversi esperimenti.
Ogni volta avrò un numero di conteggi n diverso.
La variabilità si trova sul numeratore della formula.

Regola di propagazione delle incertezze: somma in quadratura

$$\sigma_\varepsilon^2 = \left(\frac{\partial \varepsilon}{\partial n}\right)^2 \sigma_n^2 = \left(\frac{1}{N}\right)^2 \sigma_n^2 \Longrightarrow \sigma_\varepsilon = \frac{\sigma_n}{N}$$

Per una sola misura, la propagazione lineare coincide con quella in quadratura.

Efficienza di Conteggio (III)

In un esperimento di conteggio (distribuzione binomiale) si ha $\sigma_k = \sqrt{npq}$ con k variabile casuale (in questo caso k=n) e n il numero di prove (in questo caso n=N)

$$\sigma_n = \sqrt{Npq} = \sqrt{Np(1-p)} = \sqrt{N\varepsilon(1-\varepsilon)}$$

$$\sigma_\varepsilon = \frac{\sigma_n}{N} = \frac{\sqrt{N\varepsilon(1-\varepsilon)}}{N} \Longrightarrow \boxed{\sigma_\varepsilon = \sqrt{\frac{\varepsilon(1-\varepsilon)}{N}}}$$

Errore sull'efficienza per un processo binomiale

L'efficienza dipende dal rivelatore, ma se aumenta il campione diminuisce l'incertezza sull'efficienza.

Efficienza di Conteggio (IV)

Sia M il numero «vero» di particelle che hanno attraversato il rivelatore.

$$M = \frac{n}{\varepsilon}$$

La variabilità in questo caso si trova sul denominatore della formula.

Regola di propagazione delle incertezze: somma in quadratura

$$\sigma_M^2 = \left(\frac{\partial M}{\partial \varepsilon}\right)^2 \sigma_\varepsilon^2 = \left(-\frac{n}{\varepsilon^2}\right)^2 \sigma_\varepsilon^2 \Longrightarrow \sigma_M = \frac{n}{\varepsilon^2}\sigma_\varepsilon \Longrightarrow \boxed{\sigma_M = \frac{M}{\varepsilon}\sigma_\varepsilon}$$

Esercizio 6

Abbiamo un mazzo con 4 chiavi e solo una di queste apre una certa porta. Proviamo a caso le chiavi, mettendo da parte quelle che non funzionano. Definiamo la variabile casuale k = "numero di tentativi per aprire la porta", con k = 1, 2, 3, 4.

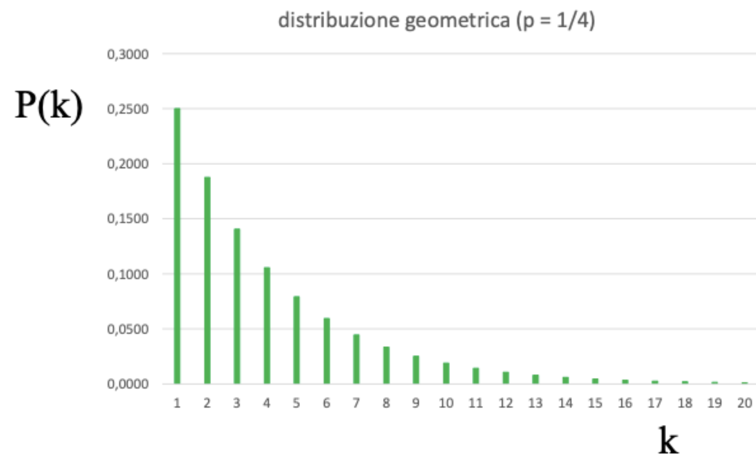
1) Qual è la distribuzione di probabilità P(k) ?
2) Quanto valgono valore medio e deviazione standard di k ?

Come cambia la distribuzione di probabilità se le chiavi provate non vengono messe da parte ?

Distribuzione Geometrica

Se le chiavi nel problema precedente non sono messe da parte, si ha la **distribuzione geometrica**.

La variabile casuale k rappresenta **il primo tentativo per il quale si verifica il "successo"** (in un processo di tipo «bernoulliano», con probabilità di successo uguale a p)



$$\boxed{P(k) = \underbrace{p}_{\text{Successo al k-esimo tentativo}}\underbrace{(1-p)^{k-1}}_{\text{Primi k-1 insuccessi}}}$$

con $k = 1 \ldots \infty$

$$\bar{k} = \frac{1}{p}$$ Media

$$\sigma_k = \frac{\sqrt{1-p}}{p}$$ Dev. St.

Distribuzione Geometrica: Normalizzazione

$$P(k) = p(1-p)^{k-1}$$

Dimostrazione della normalizzazione: $\sum_{k=0}^{+\infty} P(k) = 1$

$$\sum_{k=0}^{+\infty} \boxed{p}(1-p)^{k-1} = p \boxed{\sum_{k=0}^{+\infty}(1-p)^{k-1}} = p\frac{1}{\cancel{1}-(\cancel{1}-p)}$$

$$= p\frac{1}{p} = 1 \quad \text{cvd}$$

Coverge a $\frac{1}{1-(1-p)}$

Distribuzione Geometrica: Media

$$P(k) = p(1-p)^{k-1}$$

Dimostrazione della media: $\bar{k} = \dfrac{1}{p}$

$$\bar{k} = \sum_{k=0}^{+\infty} kP(k) = \sum_{k=0}^{+\infty} k\boxed{p}(1-p)^{k-1} = p\boxed{\sum_{k=0}^{+\infty} k(1-p)^{k-1}}$$

$$= p\frac{1}{[\cancel{1}-(\cancel{1}-p)]^2} = p\frac{1}{p^2} = \frac{1}{p} \quad \text{cvd}$$

Coverge a $\frac{1}{[1-(1-p)]^2}$

Distribuzione Geometrica: Varianza

$$P(k) = p(1-p)^{k-1}$$

Dimostrazione della varianza: $\sigma_k = \frac{\sqrt{1-p}}{p}$

$\sigma^2 = \overline{k^2} - \bar{k}^2$ dove $\bar{k} = \frac{1}{p}$ da dimostrazione precedente

$$\overline{k^2} = \sum_{k=0}^{n} k^2 P(k) = \sum_{k=0}^{n} k^2 \boxed{p}(1-p)^{k-1} = p \sum_{k=1}^{n} k^2 (1-p)^{k-1} =$$

$k^2 = k^2 - k + k$
$= k(k-1) + k$

$$= p \sum_{k=1}^{n} [k(k-1) + k](1-p)^{k-1}$$ Sommatoria di somma = somma di sommatorie

$$= p \sum_{k=1}^{n} k(k-1)(1-p)^{k-1} + \boxed{p \sum_{k=1}^{n} k(1-p)^{k-1}} = \bar{k} = \frac{1}{p}$$

Distribuzione Geometrica: Varianza (II)

$$\overline{k^2} = p\sum_{k=1}^{n} k(k-1)(1-p)^{k-1} + \frac{1}{p}$$

$$= p\sum_{k=1}^{n} \boxed{(1-p)}k(k-1)(1-p)^{k-2} + \frac{1}{p} =$$

$$= p(1-p)\boxed{\sum_{k=1}^{n} k(k-1)(1-p)^{k-2}} + \frac{1}{p} = p(1-p)\frac{2}{p^3} + \frac{1}{p} =$$

Coverge a $\frac{2}{p^3}$

$$= (1-p)\frac{2}{p^2} + \frac{1}{p}$$

$$\sigma^2 = \overline{k^2} - \bar{k}^2 = (1-p)\frac{2}{p^2} + \frac{1}{p} - \left(\frac{1}{p}\right)^2 = \frac{2-2p+p-1}{p^2} = \frac{1-p}{p^2} \qquad \text{cvd}$$

Distribuzione Geometrica (II)

La distribuzione geometrica è l'unica distribuzione di probabilità discreta caratterizzata da **assenza di memoria** descrive cioè fenomeni la cui **evoluzione futura non dipende dalla storia passata** ma solo dallo stato presente.

$$P(k+m|m\ primi\ insuccessi) = \frac{P[(k+m)\cap(m\ primi\ insuccessi)]}{P(m\ primi\ insuccessi)} =$$

Probabilità condizionata $P(A|B) = \frac{P(A\cap B)}{P(B)}$

$$= \frac{p(1-p)^{k+m-1}}{(1-p)^m} = p(1-p)^{k-1} = P(k)$$

Distribuzione Esponenziale

Altro esempio importante di assenza di memoria si ha, per variabili continue come il tempo, nel caso della **distribuzione esponenziale**

La probabilità che una particella instabile decada dopo l'istante t_0 è uguale alla probabilità che decada dopo l'istante di $(t_1 + t_0)$ se la particella non è ancora decaduta all'istante t_1.

$$\boldsymbol{P(t > t_0) = P(t > (t_1 + t_0) | t > t_1)}$$

probabilità che una particella instabile decada dopo l'istante t_0

probabilità che decada dopo l'istante di $(t_1 + t_0)$

particella non è ancora decaduta all'istante t_1

Distribuzione Esponenziale (II)

$$P(t > t_0) = P(t > (t_1 + t_0) | t > t_1)$$

Dimostrazione

$$P(t > t_0) = \int_{t_0}^{\infty} \frac{1}{\tau} e^{-\frac{t}{\tau}} dt = \left| -e^{-\frac{t}{\tau}} \right|_{t_0}^{\infty} = 0 - \left(-e^{-\frac{t_0}{\tau}}\right) = e^{-\frac{t_0}{\tau}}$$

$$P(t > (t_1 + t_0) | t > t_1) = \frac{P(t > (t_1 + t_0) \cap t > t_1)}{P(t > t_1)} = \frac{P(t > (t_1 + t_0))}{P(t > t_1)}$$

$$= \frac{\int_{t_1+t_0}^{\infty} \frac{1}{\tau} e^{-\frac{t}{\tau}} dt}{\int_{t_1}^{\infty} \frac{1}{\tau} e^{-\frac{t}{\tau}} dt} = \frac{\left| -e^{-\frac{t}{\tau}} \right|_{t_1+t_0}^{\infty}}{\left| -e^{-\frac{t}{\tau}} \right|_{t_1}^{\infty}} = \frac{0 - \left(-e^{-\frac{t_1+t_0}{\tau}}\right)}{0 - \left(-e^{-\frac{t_1}{\tau}}\right)} = \frac{e^{-\frac{t_1+t_0}{\tau}}}{e^{-\frac{t_1}{\tau}}} = e^{-\frac{t_0}{\tau}} \qquad \text{Cvd}$$

Esercizio 7

In una industria, si produce un certo tipo di elettrodomestico; il 4% dei pezzi risulta essere difettoso.

1) Se vengono selezionati, a caso, 8 pezzi, qual è la probabilità che nessuno sia difettoso ?

2) Il valore del 4% era stato ottenuto analizzando 200 pezzi e trovandone 8 difettosi. Qual è l'incertezza (pari ad una deviazione standard) da attribuire a tale probabilità ?

Riprendiamo il problema della frazione di elettrodomestici con difetti (nell'esempio trattato erano 8 su 200 pezzi esaminati). Nel caso si fosse trovato nessun pezzo difettoso su 200 esaminati, come si sarebbe potuto esprimere il valore della frazione di pezzi difettosi con un intervallo di confidenza al 95% C.L. (oppure al 99% C.L.) ?

Esercizio 8

Indichiamo con r la frazione di galassie con forma a spirale, rispetto al totale.
Consideriamo r come una variabile casuale continua, compresa tra 0 e 1.
A priori di ogni osservazione, ipotizziamo che r sia distribuita in modo uniforme tra 0 e 1.
Osserviamo poi 10 galassie, scelte a caso, e troviamo che 4 di queste hanno forma a spirale.
Qual è, in base alla formula di Bayes, la densità di probabilità della variabile r, a posteriori dell'osservazione ?

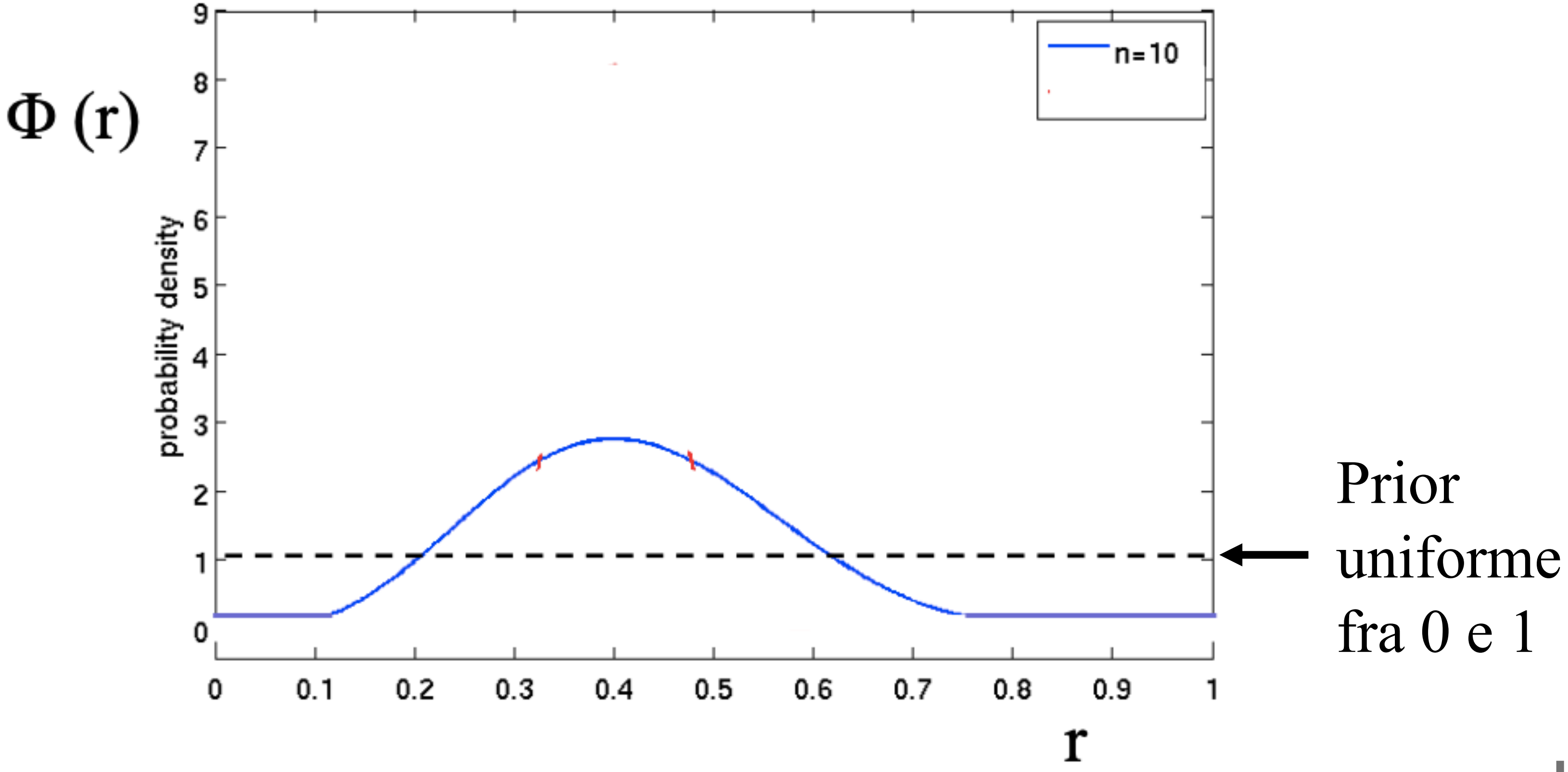
La densità di probabilità sarebbe la stessa se l'osservazione fosse stata fatta su 100 galassie, trovandone 40 a spirale ?

Formula di Bayes

$$P(H_i|E) = \frac{P(H_i)P(E|H_i)}{P(E)}$$

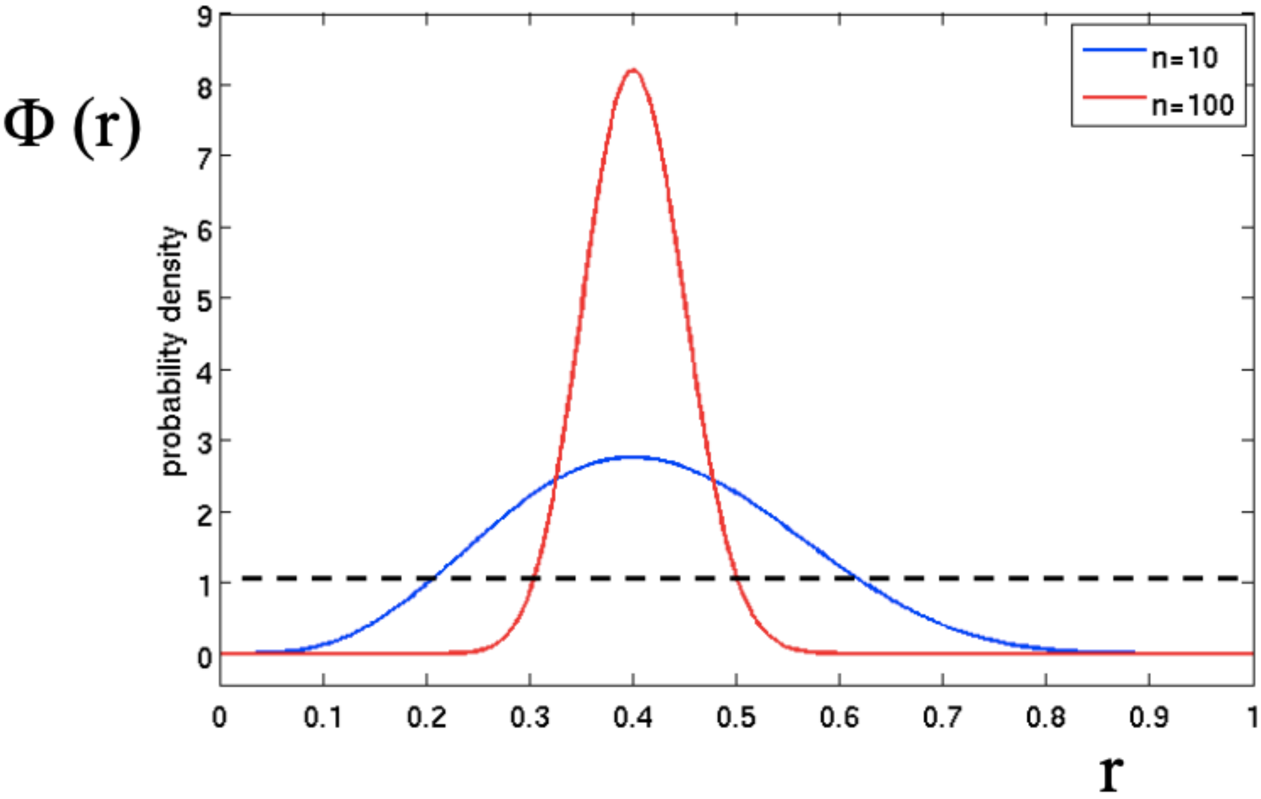
Esercizio 8 (II)

$$\Phi_{post}(r) = \binom{10}{4} r^4 (1-r)^6$$



Esercizio 8 (III)

$$\Phi_{post}(r) = \binom{100}{40} r^{40}(1-r)^{60}$$



Picco più marcato su 0.4

Distribuzione di Poisson

Immaginiamo di dovere risolvere il problema seguente

In una zona pianeggiante, omogenea, di 10 km^2 sono distribuiti, in modo casuale, 40000 alberi.
Se analizziamo una regione limitata di 1000 m^2 , qual è la probabilità di trovarvi k alberi ($0 \leq k \leq 40000$) ?

La risposta è data dalla distribuzione binomiale: il processo è bernoulliano con $p = 10^{-4}$ ed $n = 4x10^4$.
Il calcolo della probabilità risulta tuttavia piuttosto difficoltoso, dati i numeri.

Si può arrivare molto più facilmente alla soluzione grazie alla **distribuzione di Poisson**.

Distribuzione di Poisson (II)

Consideriamo un intervallo ΔX di una certa variabile (ad esempio spaziale o temporale) ed n oggetti o eventi distribuiti casualmente in ΔX.

Prendendo un sotto-intervallo $\Delta x \ll \Delta X$ la probabilità di trovarvi uno qualsiasi degli n oggetti è

$$\boldsymbol{p = \frac{\Delta x}{\Delta X} \ll 1}.$$

Se n è sufficientemente grande, ha senso definire la **densità media $\boldsymbol{\lambda}$** di oggetti per unità di spazio (o di tempo) come

$$\boldsymbol{\lambda = \frac{n}{\Delta X}}$$

Distribuzione di Poisson (III)

Il numero medio di oggetti in Δx può quindi essere espresso in termini della densità λ e della estensione Δx, senza necessariamente conoscere ΔX, n e/o p.

$$\bar{k} = \boxed{np} = (\lambda \Delta X)\left(\frac{\Delta x}{\Delta X}\right) = \lambda \Delta x = \mu$$

Media di una
distribuzione binomiale

$$\boldsymbol{\mu = \lambda \Delta x = np}$$

In pratica, in molti problemi, si conoscono λ e Δx oppure μ ma non il numero totale di "oggetti" oppure la probabilità di "successo".

Distribuzione di Poisson (IV)

La probabilità del verificarsi di un **numero k di eventi all'interno di Δx, se il numero medio per quell'intervallo è μ** (e se gli eventi sono indipendenti uno dall'altro), è data dalla

distribuzione di Poisson

$$\boxed{P(k;\mu) = \frac{\mu^k e^{-\mu}}{k!}}$$

Con k che può variare da 0 a ∞

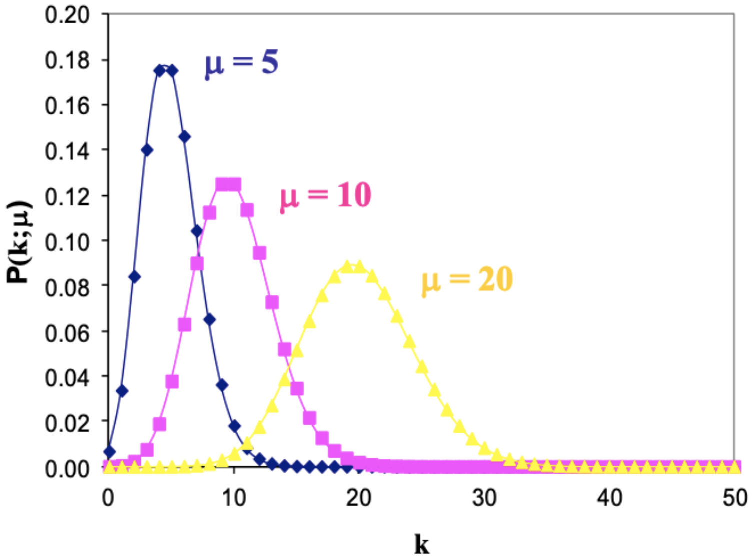
Nell'esempio precedente, $\mu = \bar{k} = 4$, da cui:

P(k=0) = 1.8%,
P(1) = 7.3%,
P(2) = 14.6%,
P(3) = 19.5%,
P(4) = 19.5%,
P(5) = 15.5%

Distribuzione di Poisson (V)

La distribuzione di Poisson è un **caso limite di binomiale**, quando **n è "grande" e p è "piccolo"** ma il valore medio, dato dal prodotto **np**, si mantiene **"finito"** (e corrisponde al parametro μ).

$$P(k;n,p) = \binom{n}{k} p^k q^{n-k} \xrightarrow[\quad]{\substack{n \to \infty \\ p \to 0 \\ np = \mu}} P(k;\mu) = \frac{\mu^k e^{-\mu}}{k!}$$



Al crescere di μ ($\boldsymbol{\mu \to \infty}$) la distribuzione si sposta a k più grandi, si allarga e diventa più simmetrica, **tendendo ad una gaussiana.**

NB: stiamo trattando una variabile discreta quindi le linee sono solo per guidare l'occhio, non può esistere una probabilità per k=0.5

Distribuzione di Poisson (VI)

La Poissoniana è il limite della binomiale. Dimostrazione:

$$P(k;n,p) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!\boxed{(n-k)!}} p^k q^{n-k} =$$

Sostituisco

$\mu = np \Longrightarrow p = \frac{\mu}{n}$

$$= \frac{\boxed{n(n-1)(n-2)\ldots(n-k+1)}}{k!} \left(\frac{\mu}{n}\right)^k \left(1-\frac{\mu}{n}\right)^{n-k}$$

$$= \frac{\mu^k}{k!} \boxed{\frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}} \boxed{\left(1-\frac{\mu}{n}\right)^{-k}} \boxed{\left(1-\frac{\mu}{n}\right)^{n}}$$

Se $n \to \infty$: $\longrightarrow 1$

Se $n \to \infty$: $\longrightarrow 1$

Limite di Eulero:
Se $n \to \infty \Longrightarrow$
$\left(1-\frac{\mu}{n}\right)^n \to e^{-\mu}$

$$P(k;n,p) = \binom{n}{k} p^k q^{n-k} \xrightarrow[\substack{p \to 0 \\ np = \mu}]{\text{Se } n \to \infty} P(k;\mu) = \frac{\mu^k e^{-\mu}}{k!}$$

Cvd

Binomiale

Poissoniana

Distribuzione di Poisson (V)

$$P(k;\mu) = \frac{\mu^k e^{-\mu}}{k!}$$ È normalizzata a 1

Dimostrazione:

$$\sum_{k=0}^{\infty} \frac{\mu^k \boxed{e^{-\mu}}}{k!} = e^{-\mu} \underbrace{\boxed{\sum_{k=0}^{\infty} \frac{\mu^k}{k!}}}_{= e^{+\mu}} = e^{-\mu} e^{+\mu} = 1$$ Cvd

Distribuzione di Poisson (VI)

$$P(k;\mu) = \frac{\mu^k e^{-\mu}}{k!}$$ Ha valor medio $\boldsymbol{\bar{k} = \mu}$

Dimostrazione:

$$\bar{k} = \sum_{k=0}^{\infty} kP(k;\mu) = \sum_{k=0}^{\infty} k\frac{\mu^k e^{-\mu}}{k!} = \sum_{k=1}^{\infty} k\frac{\mu^k e^{-\mu}}{k!} = \sum_{k=1}^{\infty} \cancel{k}\frac{\mu^{k-1}\boxed{\mu e^{-\mu}}}{\cancel{k}!}$$

$(k-1)!$

Se k=0 il termine è nullo

$\Longrightarrow$ parto da k=1

$$= \mu e^{-\mu}\sum_{k=1}^{\infty}\frac{\mu^{k-1}}{(k-1)!}$$

Cambio variabile: $h = k - 1 \rightarrow k = h + 1$

Se $k = 1 \Longrightarrow h = 0$

$$= \mu e^{-\mu}\underbrace{\boxed{\sum_{h=0}^{\infty}\frac{\mu^h}{h!}}}_{= e^{+\mu}} = \mu e^{-\mu}e^{+\mu} = \mu$$ Cvd

Distribuzione di Poisson (VII)

$$P(k;\mu)=\frac{\mu^k e^{-\mu}}{k!}$$ Ha varianza $\boldsymbol{\sigma^2=\mu}$

Dimostrazione: $\sigma^2=\overline{k^2}-\bar{k}^2$

$$\overline{k^2}=\sum_{k=0}^{\infty}k^2P(k;\mu)=\sum_{k=0}^{\infty}k^2\frac{\mu^k e^{-\mu}}{k!}=\sum_{k=1}^{\infty}k\cancel{k}\frac{\mu^k e^{-\mu}}{\cancel{k!}}=\sum_{k=1}^{\infty}k\frac{\mu^{k-1}\boxed{\mu}e^{-\mu}}{(k-1)!}$$

(k-1)!

Se k=0 il termine è nullo
$\Longrightarrow$ parto da k=1

$$=\mu\sum_{k=1}^{\infty}k\frac{\mu^{k-1}e^{-\mu}}{(k-1)!}$$

Cambio variabile: $h=k-1\rightarrow k=h+1$
Se $k=1\Longrightarrow h=0$

$$=\mu\sum_{h=0}^{\infty}(h+1)\frac{\mu^h e^{-\mu}}{h!}=\mu\boxed{\sum_{h=0}^{\infty}h\frac{\mu^h e^{-\mu}}{h!}}+\mu\boxed{\sum_{h=0}^{\infty}\frac{\mu^h e^{-\mu}}{h!}}$$

$= 1$ (normalizzazione)

$=\mu$ (media)

Distribuzione di Poisson (VIII)

$$\overline{k^2} = \mu\mu + \mu 1 = \mu^2 + \mu$$

$$\sigma^2 = \overline{k^2} - \bar{k}^2 = (\mu^2 + \mu) - \mu^2 = \mu$$ Cvd

Allo stesso risultato si arriva facendo il limite della varianza binomiale

$$\sigma^2 = npq = np(1-p) = np - npp \xrightarrow[\substack{n \to \infty \\ p \to 0 \\ np = \mu}]{} \mu - \mu 0 = \mu$$ Cvd

Esercizio 9

Prepariamo biscotti con gocce di cioccolato.
Mettiamo nell'impasto 300 gocce e facciamo 100 biscotti.

1) Qual è la probabilità che un biscotto, preso a caso, non contenga nessuna goccia ?

2) Quante gocce di cioccolato dovremmo mettere nello stesso impasto (per 100 biscotti) affinché tale probabilità risulti dell'1% ?

Esercizio 10

Ipotizziamo che il protone sia una particella instabile e decada in modo esponenziale, con vita media τ (non nota ma, sicuramente, "molto grande"):

$$N(t) = N_0 e^{-\frac{t}{\tau}}$$

In un esperimento, si osserva un numero N_0 di protoni per un tempo $\Delta t << \tau$ e non si registra alcun decadimento.

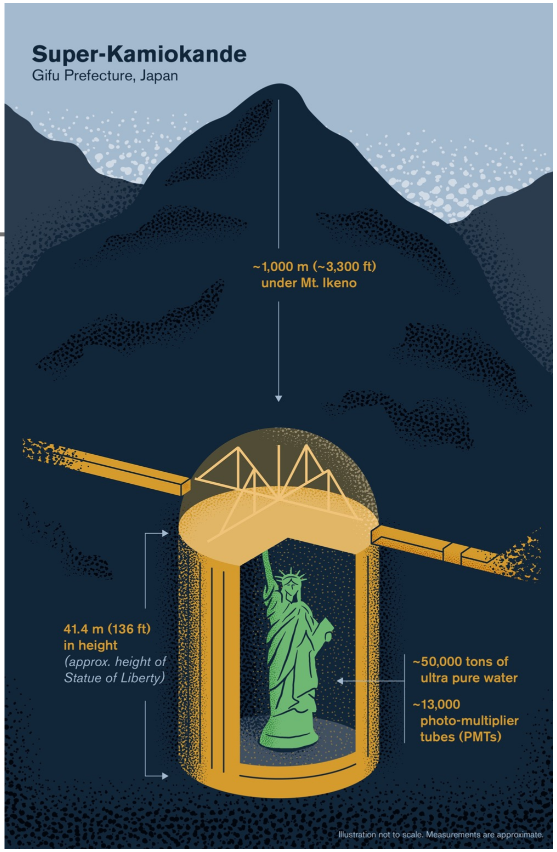
1) Qual è la probabilità P(0) di osservare nessun decadimento in un tempo Δt, se la vita media è τ (cioè, in funzione di τ) ?

2) In base al risultato precedente, quale limite inferiore si può dare a τ, fissato un determinato livello di confidenza ?

Vita Media del Protone

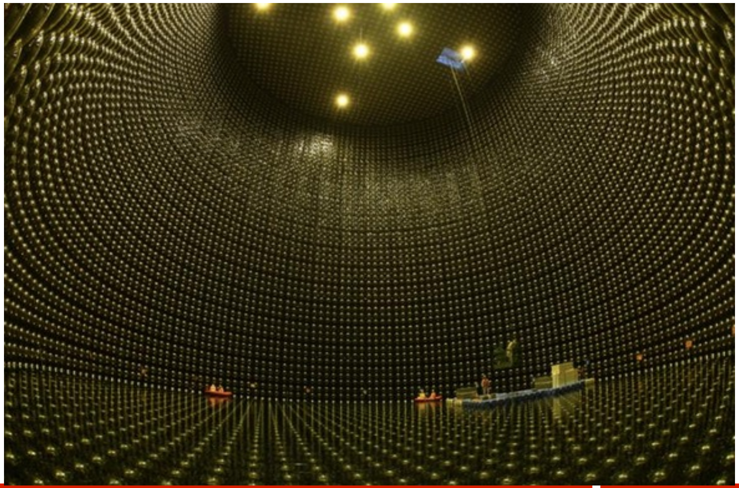
L'esperimento **Super Kamiokande** ha osservato $N_0 \approx 10^{33}$ protoni per $\Delta t \approx 10$ anni, cercando il decadimento

$$p \to e^+\pi^0 \to e^+\gamma\gamma$$



Proibito nel Modello Standard ma ipotizzato nelle Teorie di Grande Unificazione (GUT).

Esperimento di conteggio → Limite inferiore sulla vita media del protone.



Vita media protone $\tau > 10^{30}$ anni
Età universo 10^{10} anni

La materia è stabile!!

Distribuzione di Poisson (IX)

Quando si effettua una **singola misura** di conteggio, diciamo m, di una variabile poissoniana:

- **la miglior stima del parametro** μ è la misura stessa

$$\hat{\mu} = \mathrm{m}$$

- **l'incertezza** da associare alla misura è la radice quadrata della misura stessa

$$\Delta_m = \sqrt{m}$$

- L'errore relativo quindi decresce all'aumentare di m

$$\frac{\Delta m}{m} = \frac{\sqrt{m}}{m} = \frac{1}{\sqrt{m}}$$

Distribuzione di Poisson (X)

La migliore stima del numero medio di conteggi attesi è pari alla media matematica $\hat{\mu} = \bar{m}$

<u>Dimostrazione:</u>
Uso il criterio di massima verosimiglianza

Supponiamo di aver un campione di misure $\{m_{1,} m_2, \ldots m_N\}$ estratte da una popolazione poissoniana

$$P(m_i) = \frac{\mu^{m_i} e^{-\mu}}{m_i!}$$

La probabilità di ottenere l'intero campione è la probabilità combinata delle diverse misure (indipendenti fra loro)

$$P(campione) = \prod_{i=1}^{N} \frac{\mu^{m_i} e^{-\mu}}{m_i!} = \frac{\mu^{m_1} e^{-\mu}}{m_1!} \cdot \frac{\mu^{m_2} e^{-\mu}}{m_2!} \cdot \ldots \frac{\mu^{m_N} e^{-\mu}}{m_N!} = \frac{\mu^{\sum_{i=1}^{N} m_i} e^{-N\mu}}{m_1!\, m_2! \ldots m_N!}$$

NB: la variabile casuale è μ finchè non massimizzo la probabilità

Distribuzione di Poisson (XI)

Considero il logaritmo della funzione per passare dal prodotto alle somme. Il logaritmo è un funzione crescente nell'argomento, quindi non cambia il valore massimo.

$$lnP(campione) = ln\left(\frac{\mu^{\sum_{i=1}^{N} m_i} e^{-N\mu}}{m_1!\, m_2! \dots m_N!}\right)$$

$$= ln\left(\mu^{\sum_{i=1}^{N} m_i} e^{-N\mu}\right) - ln(m_1!\, m_2! \dots m_N!) =$$

$$= ln\left(\mu^{\sum_{i=1}^{N} m_i}\right) + ln(e^{-N\mu}) - ln(m_1!\, m_2! \dots m_N!) =$$

$$= \left(\sum_{i=1}^{N} m_i\right) ln(\mu) - N\mu\, lne - ln(m_1!\, m_2! \dots m_N!)$$

$$= \left(\sum_{i=1}^{N} m_i\right) ln(\mu) - N\mu + cost$$

Distribuzione di Poisson (XII)

Il massimo si ha quando la derivata si annulla

$$\frac{\partial P}{\partial \mu} = 0 \Longrightarrow \frac{\partial\left(\left(\sum_{i=1}^{N} m_i\right) ln(\mu) - N\mu + cost\right)}{\partial \mu} = 0 \Longrightarrow$$

$$\frac{\partial\left(\boxed{\left(\sum_{i=1}^{N} m_i\right)} ln(\mu)\right)}{\partial \mu} - \frac{\partial(\boxed{N}\mu)}{\partial \mu} + \cancel{\frac{\partial(cost)}{\partial \mu}} = 0 \Longrightarrow$$

$$\sum_{i=1}^{N} m_i \frac{1}{\mu} - N = 0 \Longrightarrow \hat{\mu} = \frac{\sum_{i=1}^{N} m_i}{N} = \overline{m}$$

Cvd

Esercizio 11

Sappiamo che il tasso medio di decadimento di un campione radioattivo è di circa 20 disintegrazioni al minuto.
Quanto a lungo deve durare una misura affinché l'errore relativo sulla misura del tasso di decadimento sia dell'ordine del 10% ?
E dell' 1% ?

Esercizio 12

Un certo tipo di batterio è presente, in aria, con una densità media di 100 batteri per metro cubo.
Qual è la probabilità che, all'interno di un certo volume ampio $V = 2\ dm^3$ e scelto a caso, sia presente almeno un batterio ?

Esercizio 13

In un dato processo, il numero di conteggi segue la distribuzione di Poisson. Sapendo che la probabilità di osservare zero conteggi è il 13.5%, qual è la probabilità di osservare almeno tre conteggi?

Esercizio 14

Nella fabbricazione di tappeti per pavimenti, si verificano piccoli difetti distribuiti casualmente con un tasso medio di 0.96 per 20 m^2 di superficie. I pavimenti degli uffici di un certo edificio sono stati tappezzati con questo materiale. Qual è la probabilità che nel pavimento di un dato ufficio, la cui superficie è di 50 m^2, siano presenti al massimo tre difetti ?

Esercizio 15

Utilizzando un opportuno sistema ottico di misura si conta il numero di sferette micrometriche distribuite casualmente e indipendentemente su una grande superficie. In media, si misurano 576 sferette al centimetro quadro.

1) Qual è l'area massima di osservazione per la quale $P(k=0)$ è maggiore dell'1% ? Il simbolo k indica il numero di sferette presenti all'interno dell'area.
2) Qual è la probabilità $P(k>2)$ in corrispondenza dell'area massima determinata nella soluzione al punto 1) ?

Esercizio 16 – Backup

Si ha a disposizione per T = 30 min un contatore Geiger per misurare il tasso di decadimento R_S di una sorgente radioattiva.
Si misura R_S come la differenza $R_S = R_{S+F} - R_F$ tra il tasso della sorgente+fondo ambientale e il tasso del solo fondo:

$$R_S = R_{S+F} - R_F = (N_{S+F} / T_{S+F}) - (N_F / T_F)$$

dove N_i = numero di conteggi e T_i = tempo di conteggio (avente errore trascurabile), con il vincolo che $T_{S+F} + T_F = T$.

Sapendo, da precedenti misure, che $R_F/R_{S+F} \approx 1/16$, quanto deve durare T_F per minimizzare l'errore ΔR_S ?
(sommare le incertezze in quadratura)