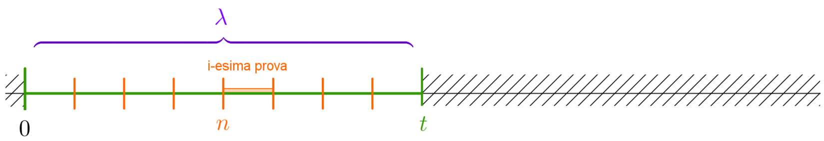
2.2 Il caso generale

Torniamo a considerare il problema generale di conteggio descritto nel paragrafo 1.2. Un possibile modello probabilistico per descrivere la situazione è quello binomiale, quindi proviamo a modellizzare la situazione mediante quest'ultima distribuzione, in modo analogo a quanto visto nell'esempio del centralino.

2.2.1 Costruzione del modello

L'idea
Suddividiamo l'intervallo $[0, t]$ in ***n*** **intervallini uguali** in modo che in ogni intervallino si abbia "al più" una sola realizzazione dell'evento.



La situazione si può interpretare nel modo seguente.

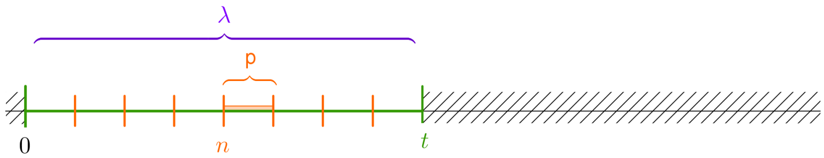
- Consideriamo una **sequenza**[6] di n **prove**.
- Ciascuna prova ha **due soli esiti** possibili: "*l'evento si realizza*", "*l'evento non si realizza*" nell'intervallino ad essa relativo[7].

  Siccome vogliamo al più un solo successo su ogni intervallino, consideriamo ***n*** **"grande"** rispetto al numero medio di successi nell'intervallo $[0, t]$.

- A questo punto dobbiamo determinare la **probabilità** p che l'evento si realizzi nell'intervallino. L'idea è di utilizzare un approccio frequentista, ossia considerare la frequenza relativa dell'evento in esame sull'insieme delle n prove.
  Abbiamo così:
  $$p = \frac{\#\ \textit{realizzazioni evento}}{\#\ \textit{prove}} = \frac{\lambda}{n}$$

  Assumiamo che ***p*** sia **costante** per ogni prova. Osserviamo che ciò equivale ad ipotizzare che il numero medio di realizzazioni sia costante su tutti gli intervallini considerati.

- Inoltre le prove sono **indipendenti**, ovvero nessuna realizzazione è influenzata dalle altre.



Il modello ottenuto, quindi, è la **distribuzione binomiale** con parametri n e p definiti sopra.

[6]Tale sequenza si indica spesso come "schema di Bernoulli" o "schema successo-insuccesso".

[7]Anche in questo caso spesso l'esito che corrisponde alla realizzazione dell'evento si indica come "*successo*"; l'altro come "*insuccesso*". Nell'esempio del centralino, considerato in precedenza, l'evento è l'arrivo di una telefonata.

2.2.2 In sintesi

Abbiamo così risolto la questione di conteggio posta all'inizio, infatti abbiamo fornito un modello: la **distribuzione binomiale** che ha parametri n e p del tipo indicato (in particolare tali che $np = \lambda$).
Di tale modello è nota l'espressione analitica. Precisamente:

* sia S la **variabile aleatoria** che conta il numero di realizzazioni dell'evento nell'intervallo $[0, t]$;
* la **probabilità** che S assuma valore k è

$$\boldsymbol{P(S = k) = \binom{n}{k} \cdot p^n \cdot (1-p)^{n-k}}, \qquad \text{per } k = 0, 1, 2, ...$$

Dove con il simbolo $\binom{n}{k}$ si denota il coefficiente binomiale n su k, la cui espressione esplicita è $\dfrac{n!}{k!(n-k)!}$ [8].

2.2.3 Valutazione del modello

Il modello binomiale così costruito:

- **non** è **efficiente** dal punto di vista computazionale
  visto che la probabilità è espressa in termini dei coefficienti binomiali e il loro calcolo è articolato, infatti, il fattoriale è una funzione che cresce molto rapidamente:

$$10! \simeq 3.6 \cdot 10^6, \qquad 20! \simeq 2,4 \cdot 10^{18}, \qquad 70! > 10^{100}$$

- **non** è molto **espressivo**
  infatti, il parametro λ, che caratterizza la situazione, non compare esplicitamente.

Quindi è utile costruire un nuovo modello per rappresentare la situazione.

[8] Dato un numero naturale m, si indica con $m!$ il prodotto $m \cdot (m-1) \cdot ... \cdot 2 \cdot 1$, se $m \geq 1$; e 1 se $m = 0$.