## Probabilità

- **Teo. Prob. Totali**: $P(\bigcup_{i=1}^{n} P(A_i)) = \sum_{i=1}^{n} P(A_i)$; $A_1, ..., A_n$ eventi indipendenti.
- **Prob. Condizionate**: $P(A|B) = P(A \cap B)/P(B)$.
- **Teo. Prob. Totali (Partizioni)**: $P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$, $B_1, ..., B_n$ partizioni di $\Omega$.
- **Teo. Bayes**: $P(A_i|B) = P(B|A_i)P(A_i)/(\sum_{j=1}^{n} P(B|A_j)P(A_j) = P(B))$, $A_i \in A_1, ..., A_n$ partizioni di $\Omega$.
- **Note sui Complementari**: $P(A) = 1 - P(A^C)$; $P(A|B) = 1 - P(A^C|B)$.
- **Indipendenza**: $A_1, ..., A_n$ eventi indipendenti sse $P(\bigcap_{i=1}^{n} P(A_i)) = \prod_{i=1}^{n} P(A_i)$.
- **Note sull'Indipendenza**: $P(A|B) = P(A)$ sse $A, B$ eventi indipendenti; indipendenza condizionata $\not\Leftrightarrow$ indipendenza incondizionata.

## Calcolo Combinatorio

- **Permutazioni**: $n$ elementi distinti $\rightarrow n!$ permutazioni; se $m \leq n$ elementi sono indistinguibili $\rightarrow n!/m!$ disposizioni.
- **Sottoinsiemi di $k$ elementi**: $n$ elementi, assunto $k \leq n \rightarrow n!/(n-k)!$ sottoinsiemi; se l'ordine degli elementi all'interno dei sottoinsiemi non mi importa avrò $\binom{n}{k} = n!/k!(n-k)!$ sottoinsiemi.
- **Prob. Binomiale**: dati $n$ tentativi, $P(succ.) = p$, assunto $k \leq n \rightarrow P(k\ succ./n\ tentativi) = \binom{n}{k} p^k (1-p)^{n-k}$.
- **Partizioni**: dato $\Omega$ partizionato in $K_n$ sottoinsiemi, calcolo la probabilità, su $n$ tentativi, di ottenere $k_1 \in K_1, ..., k_n \in K_n$ elementi $\rightarrow \binom{n}{k_1, ..., k_n} = n!/k_1!k_2!...k_n$.
- **Prob. Ipergeometrica**: dato $|\Omega| = n$ partizionato in $K_1 = k$ $K_2 = n - k$ sottoinsiemi e scelto un campione di $c < n$ elementi voglio calcolare la probabilità che questo sia composto da $k' \leq k \in K_1$ e $k'' = c - k' \leq n - k \in K_2$ elementi $\rightarrow \binom{k}{k'}\binom{n-k}{c-k'}/\binom{n}{c}$.

## Variabili Aleatorie Discrete

- **Legge di Prob.**: $p_X(x) = P(X = x)$.
- **Valore Atteso**: $E[X] = \sum_x x \cdot p_X(x)$; $E[X|A] = \sum_x x \cdot p_{X|A}(x)$ se condizionato.
- **Legge dello Statistico Inconsapevole**: sia $g(x)$ deterministica $\rightarrow E[g(X)] = \sum_x g(x) \cdot p_X(x)$.
- **Linearità del Valore Atteso**: $E[\alpha X + \beta] = \alpha E[X] + \beta$; se $g(x)$ è determinista e lineare $E[g(X)] = g(E[X])$.
- **Varianza**: $Var[X] = E[X^2] - E[X]^2 = E[(X - E[X])^2]$.
- **Semi-Linearità della Varianza**: $Var[\alpha X + \beta] = \alpha^2 Var[X]$.
- **Deviazione Standard**: $\sigma_X = \sqrt{Var[X]}$.
- **Perdita di Memoria**: $p_{X-t|X>t}(x) = p_X(x)$.
- **Legge dell'Aspettativa Totale**: $E[X] = \sum_{i=1}^{n} P(A_i)E(X|A_i)$, $A_1, ..., A_n$ eventi che partizionano $\Omega$.

## Variabili Aleatorie Discrete Multiple

- **Marginalizzazione**: $p_Y(y) = \sum_x p_{X,Y}(x, y) = \sum_x p_X(x) \cdot p_{Y|X}(y|x)$.
- **Valore Atteso**: data $g(x, y)$ deterministica $E[g(X, Y)] = \sum_x \sum_y g(x, y) \cdot p_{X,Y}(x, y)$; caso particolare $E[X+Y] = E[X]+E[Y]$.
- **Varianza**: $Var[X+Y] = Var[X]+Var[Y]+2(E[XY]-E[X]E[Y])$.
- **Casi Particolari (X, Y Indipendenti)**:

$$X \perp Y \Rightarrow \begin{cases} E[XY] = E[X]E[Y] \\ E[g(X)h(Y)] = E[g(X)]E[h(Y)] \\ Var[\alpha X + \beta Y] = \alpha^2 Var[X] + \beta^2 Var[Y] \end{cases}$$

## Variabili Aleatorie Continue

- **Densità di Prob.**: $P(x \leq X \leq x + \delta) = \int_x^{x+\delta} f_X(\gamma)d\gamma$.
- **Valore Atteso**: $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.
- **Legge dello Statistico Inconsapevole**: $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$.
- **Varianza**: $\sigma^2 = Var[X] = E[X^2] - E[X]^2 = \int_{-\infty}^{\infty} (x - E[X])^2 \cdot f_X(x)dx$.
- **Cumulata di Prob.**: $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(\gamma)d\gamma$.

## Variabili Aleatorie Continue Multiple

- **Marginalizzazione**: $f_Y(y) = \int_x f_{X,Y}(x, y)dx = \int_y f_{Y|X}(y|x) f_X(x)dx$.
- **Valore Atteso**: data $g(x, y)$ deterministica $E[g(X, Y)] = \iint_{\mathbb{R}} g(X, Y) f_{X,Y}(x, y)dxdy$.
- **Casi Particolari (X, Y Indipendenti)**: $X \perp Y \Rightarrow f_{X,Y}(x, y) = f_X(x)f_Y(y) \wedge F_{X,Y}(x, y) = F_X(x)F_Y(y)$.
- **Teo. Bayes nel Continuo e Situazioni Ibride**:
  - $X, Y$ continue: $f_{X|Y}(x, y) = f_{Y|X}(y|x)f_X(x)/f_Y(y)$.
  - $X$ discreta, $Y$ continua: $p_{X|Y}(x|y) = f_{Y|X}(y|x)p_X(x)/f_Y(y)$.
  - $X$ continua, $Y$ discreta: $f_{X|Y}(x|y) = p_{Y|X}(y|x)f_X(x)/p_Y(y)$.
- **Somma di Variabili Aleatorie Continue**: Sia $Z = X + Y$:
  - $X \perp Y \iff f_{Z,X}(z, x) = f_X(x)f_Y(z - x)$.
  - $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z - x)dx$. **Convoluzione.**
  - $F_Z(z) = F_{X+Y}(z) = \int_{-\infty}^{+\infty} F_x(z - y)f_Y(y)dy$.

## Calcolo di una Funzione di V.A.

- **Discr.** $p_Y(y) = P(g(X) = y) = \sum_{x|g(x)=y} p_X(x)$.
- **Cont. (Cumulata)** $F_Y(y) = P(g(X) \leq y) \implies f_Y(y) = dF_y(y)/dy$.
- **Cont. ($g(X)$ Lineare)** $Y = \alpha X + \beta \implies f_Y(y) = \frac{1}{|\alpha|} f_X(\frac{x-\beta}{\alpha})$.
- **Cont. ($g(X)$ Monotona)** $Y = g(X) \implies f_Y(y) = f_X(g^{-1}(y))/\left|\frac{d}{dx}g(g^{-1}(y))\right|$.

## Altri Indicatori Statistici per V.A. Multiple

- **Covarianza**
  - $Cov[X, Y] = E[XY] - E[X]E[Y]$.
  - $Cov[X, X] = Var[X]$; $Cov[\alpha X, Y] = \alpha Cov[X, Y]$.
  - $X \perp Y \Rightarrow Cov[X, Y] = 0$ ma $Cov[X, Y] = 0 \not\Rightarrow X \perp Y$.
  - $Cov[\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j] = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov[X_i, Y_j]$.
  - $Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i] + 2\sum\sum_{i<j} Cov[X_i, X_j]$.
- **Coeff. Correlazione Lineare**: $\rho = Cov[X, Y]/\sigma_X\sigma_Y$. Se $\rho = 1 \implies (X - E[X]) = \alpha(Y - E[Y])$.
- **Valore Atteso Condizionato**: $E[X|Y] = \sum_x x \cdot p_{X|Y}(x|y)$.
- **Legge delle Aspettazioni Iterate**: $E[X] = E[E[X|Y]]$.
- **Varianza Condizionata**: $Var[X|Y] = E[X^2|Y = y] - E[X|Y = y]^2$.
- **Legge della Variazione Totale**: $Var[X] = E[Var[X|Y]] + Var[E[X|Y]]$.

## Successioni di V.A.

- **Somma di $N$ (casuale) V.A. Indipendenti $X_i$**:
  - $E[\sum_{i=1}^{N} X_i] = E[E[\sum_{i=1}^{N} X_i N]] = E[N]E[X]$.
  - $Var[X] = E[N]Var[X_1] + Var[N]E[X_1]^2$.
- **Diseguaglianza di Markov**: $E[X] \geq \alpha P(X \geq \alpha)$.
- **Diseguaglianza di Chebyshev**: $Var[X] \geq \alpha^2 P(|X - E[X]| \geq \alpha)$.
- **Convergenza in Prob.**: Sia $A_k$ una successione di V.A. e sia $\alpha \in \mathbb{R}$; $A_k$ si dice convergente in probabilità ad $\alpha$ se: $\lim_{k\to+\infty} P(|A_k - a| \geq \epsilon) = 0\ \forall\epsilon > 0$. $A_k \to^P \alpha$.
- **Media Campionaria**: Siano $X_1, X_2, .., X_n$ V.A. I.I.D.; $M_n = (X_1 + X_2 + .. + X_n)/n$. $M_n \to^P E[X]$; $E[M_n] =^{n\to+\infty} E[X]$; $Var[M_n] =^{n\to+\infty} 0$.

## V.A. Notevoli

- **V.A. Uniforme** $X \sim \mathcal{U}[\alpha, \beta]$ (cont.):
  - "Attribuisce la stessa probabilità ad ogni elemento dell'insieme $S$ su cui è definita".
  - $f_X(x) = \begin{cases} 1/(\beta - \alpha) & \alpha < x < \beta \\ 0 & altr. \end{cases}$
  - $F_X(x) = \begin{cases} 0 & x < \alpha \\ (x - \alpha)/(\beta - \alpha) & \alpha < x < \beta \\ 1 & x > \beta \end{cases}$
  - $E[X] = \frac{1}{2}(\alpha + \beta)$; $Var[X] = \frac{1}{12}(\beta - \alpha)^2$.
- **V.A. Uniforme** $X \sim \mathcal{U}\{\alpha, \beta\}$ (discr.):
  - $p_X(x) = 1/n$.
  - $E[X] = \frac{1}{2}(\alpha + \beta)$; $Var[X] = \frac{1}{12}(n^2 - 1)$.
- **V.A. Geometrica** $X \sim \mathcal{G}(p)$ (discr.):
  - "Probabilità che il primo successo (o evento in generale) richieda l'esecuzione di $x$ prove indipendenti, ognuna con probabilità di successo $p$".
  - $p_X(x) = (1 - p)^{x-1}p$.
  - $E[X] = 1/p$; $Var[X] = (1 - p)/p^2$.
- **V.A. Binomiale/di Bernoulli** $X \sim \mathcal{B}(n, p)$ **(se $n = 1$ Binomiale $\rightarrow$ Bernulliana)** (discr.):
  - "Probabilità di ottenere $i$ successi in $n$ prove indipendenti, ognuna con probabilità di successo $p$".
  - $p_{X_i}(1/succ.) = p$; $p_{X_i}(0/insucc.) = 1 - p$.
  - $p_B(i) = p_{X_1+X_2+...+X_n}(i) = \binom{n}{i}p^i(1 - p)^{n-i}$.
  - $E[X] = np$; $Var[X] = np(1 - p)$.
- **V.A. Esponenziale** $X \sim \mathcal{E}(\lambda)$ (cont.).
  - $f_X(x) = \begin{cases} \lambda e^{-x\lambda} & x \geq 0 \\ 0 & x < 0 \end{cases}$.
  - $F_X(x) = 1 - e^{-x\lambda}$.
  - $E[X] = 1/\lambda$; $Var[X] = 1/\lambda^2$.
- **V.A. Gaussiana** $X \sim \mathcal{N}(\mu, \sigma^2)$ (cont.).
  - $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.
  - $E[X] = \mu$; $Var[X] = \sigma^2$.
  - $\mathcal{N}(0, 1)$ è la gaussiana standard i cui valori di cumulata sono tabulati; $X \sim \mathcal{N}(\mu, \sigma^2) \implies \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$.
  - $Y = \alpha X + \beta \implies Y \sim \mathcal{N}(\alpha\mu + \beta, \alpha^2\sigma^2)$.
  - $X \perp Y$; $X \sim \mathcal{N}(\mu', \sigma_{'}^2)$, $Y \sim \mathcal{N}(\mu'', \sigma_{''}^2)$; $Z = X + Y \sim \mathcal{N}(\mu' + \mu'', \sigma_{'}^2 + \sigma_{''}^2)$. In ogni caso, anche se $X \not\perp Y$, $E[Z] = E[X] + E[Y]$; la varianza, invece, necessita del fattore correttivo. La somma di gaussiane sarà sempre gaussiana.

## Teorema Fondamentale del Limite

- **C.L.T.**: Siano $X_i$ v.a. i.i.d. con $Var[X_i] = \sigma^2$ finita e $E[X_i] = \mu$. Allora, $S_n = X_1 + X_2 + ... + X_n$ è gaussiana per $n \to \infty$; normalizzando: $Z_n = \frac{S_n - E[S_n]}{\sqrt{Var[S_n]}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- **Problema del sondaggista**: Sia $M_n$ una media campionaria di v.a. i.i.d. $X_i$ con $E[X_i] = \mu$ e $Var[X_i] = \sigma^2$. Allora, per $n \to \infty$, $P(|M_n - E[M_n]| < \alpha) \geq \beta$ è calcolabile come: $P\left(|Z| < \frac{\alpha\sqrt{n}}{\sigma}\right) \geq \beta$ dove $Z \sim \mathcal{N}(0, 1)$.
- **Th. De Moivre-Laplace**: Sia $X \sim \mathcal{B}(n, p)$, allora $Z_n = \frac{X - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0, 1)$ per $n \to \infty$. Inoltre, $P(a < X < b) \approx \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right)$. Se è necessario usare questo metodo per approssimare un coefficiente binomiale, $p = 0.5$ risulta comodo.

## Processi Casuali

- **Proc. di Bernoulli**: $BP(p)$ è una serie di v.a. i.i.d. $X_i \sim Bern(p)$.
  - *Numero di successi (arrivi) $S$ in $n$ istanti temporali*: $S \sim \mathcal{B}(n, p)$.
  - *Tempo di interarrivo*: $T_i \sim \mathcal{G}(p)$. Tutti i tempi di interarrivo sono indipendenti e godono di perdita di memoria.
  - *Tempo al k-esimo arrivo*: $Y_k = T_1 + T_2 + ... + T_k$.

$Y_k \sim Pascal-k(p)$ v.a. di Pascal.
    - $P(Y_k = t) = P(k-1\ arr.in\ [1, t-1], arr.\ in\ t) = p\binom{t-1}{k-1}p^{k-1}(1-p)^{t-k}$ assunto $t \geq k \geq 1$.
    - $E[Y_k] = k/p$.
    - $Var[Y_k] = k(1-p)/p^2$.
  - *Splitting di un B.P.*: gli arrivi di $BP(p)$ possono essere "divisi" in due sottoprocessi $BP(pq)$ e $BP(p(1-q))$ con $q$ prob., per un arrivo, di finire nel primo e $1-q$ di finire nel secondo b.p.; i due processi sono indipendenti.
  - *Merging di due B.P.*: gli arrivi di $BP(p)$ e $BP(q)$ possono essere riuniti in un unico $BP(p+q-pq)$.
- **Proc. di Poisson**: $PP(\lambda)$ è la versione continua dei B.P.
  - *Probabilità di avere k arrivi nell'intervallo* $[0,\tau]$: $N[0,\tau] \sim Poisson(\lambda\tau)$ è una v.a. di Poisson.
    - $P(N[0,\tau] = k) = \begin{cases} \frac{(\lambda\tau)^k}{k!}e^{-\lambda\tau} & k \in \mathbb{N} \\ 0 & altr. \end{cases}$.
    - $E[N[0,\tau]] = Var[N[0,\tau]] = \lambda\tau$.
    - $\sum_{k=0}^{+\infty} P_{N[0,\tau]}(k) = 1$.
  - *Tempo di interarrivo*: $T_i \sim \mathcal{E}(p)$. Tutti i tempi di interarrivo sono indipendenti e godono di perdita di memoria.
  - *Distribuzione del tempo al k-esimo arrivo*: $Y_k \sim Erlang-k(\lambda)$, ossia una somma di $k$ esponenziali (interarrivi).
    - $f_{Y_k}(t) = \begin{cases} \frac{(\lambda t)^{k-1}}{(k-1)!}e^{-\lambda t}\lambda & t > 0, k \geq 1 \\ 0 & altr. \end{cases}$
    - $E[Y_k] = \frac{k}{\lambda}$, $Var[Y_k] = \frac{k}{\lambda^2}$.
  - *Splitting di un P.P.*: $PP(\lambda) \to PP(\lambda q) \wedge PP(\lambda(1-q))$.
  - *Merging di due P.P.*: $PP(\lambda_1) \wedge PP(\lambda_2) \to PP(\lambda_1 + \lambda_2)$.
  - *Incidenza casuale per P.P.* Dato un P.P. iniziato da un tempo indefinito $t$, il tempo tra l'arrivo precedente a $t$ e quello successivo è una $Erlang-2(\lambda)$.
- Relazione Bernoulli/Poisson:

| | Poisson | Bernoulli |
|---|---|---|
| Tempo di arrivo | Continuo | Discreto |
| Rate degli arrivi | $\lambda$/unità di tempo | $p$/per prova |
| Ddp del numero di arrivi | $Poisson(\lambda t)$ | $Bin(n,p)$ |
| Ddp del tempo di interarrivo | $Exp(\lambda)$ | $Geom(p)$ |
| Ddp del tempo al k-esimo arrivo | *Erlang-k*$(\lambda)$ | *Pascal-k*$(p)$ |

## Stima Bayesiana

- A partire da una v.a. $X$, che rappresenta l'esito di una misura, si vuole stimare, mediante la regola di Bayes, $\Theta$ (v.a.) quantità da misurare:
  - $f_{\Theta|X}(\theta|X) = \frac{f_{X|\Theta}(x|\theta)f_\Theta(\theta)}{f_X(x)}$.
  - $f_X(x) = \int_{\mathbf{R}} f_{X|\Theta}(x|\theta')f_\Theta(\theta')d\theta'$.
- **Stimatore "M.A.P."**:
  - $\hat{\Theta}_{MAP} = \begin{cases} \arg\max(\Theta)P_{\Theta|X}(\theta|X) & discr. \\ \arg\max(\Theta)f_{\Theta|X}(\theta|X) & cont. \end{cases}$
  - Alternativamente, se $\Theta$ v.a. cont., $E[\Theta|X = x] = \int_{\mathbb{R}} \theta f_{\Theta|X}(\theta|x)d\theta$ (vedi sotto).
- **Stimatore "L.M.S"**:
  - $\hat{\Theta}_{LMS} = E[\Theta] \to \hat{\Theta}_{LMS}(X) = E[\Theta|X]$.
  - *M.S.E.*: *error*$(\hat{\Theta}_{LMS}) = E[(\Theta - \hat{\Theta}_{LMS})^2] = Var[\Theta]$.
  - Min E.Q.M.: $E[(\Theta - E[\Theta|X])^2] = E[Var[\Theta|X]] \leq Var[\Theta]$.
  - $Var[\Theta] = E[Var[\Theta|X]] + Var[E[\Theta|X]] \geq E[Var[\Theta|X]]$.
  - *Proprietà*:
    - Errore di stima: $\Theta = \hat{\Theta}_{LMS} - \widetilde{\Theta}$, $E[\widetilde{\Theta}] = 0$ (LMA apolarizzata: nè sottostimiamo, nè sovrastimiamo $\Theta$). Per qualunque $h(X)$ deterministica: $E[\widetilde{\Theta} \cdot h(X)] = 0$.
    - $Cov[\widetilde{\Theta}, \hat{\Theta}_{LMS}] = 0$.
    - $Var[\widetilde{\Theta}] = E[Var[\Theta|X]]$.
- **Stimatore "L.M.S." lineare**:
  - $\hat{\Theta}_{LIN}(X) = E[\Theta] + \frac{Cov[X,\Theta]}{Var[X]}(X - E[X])$.
  - *Standardizz.*: $\frac{\hat{\Theta}_{LIN}(X) - E[\Theta]}{\sigma_\Theta} = \frac{Cov[X,\Theta]}{\sigma_X\sigma_\Theta}\frac{X-E[X]}{\sigma_X}$.
  - *M.S.E. lineare*: $MSE = E[(\Theta - \hat{\Theta}_{LIN})^2] = (1 - \rho^2[X,\Theta])Var[\Theta] = (1 - \frac{Cov^2[X,\Theta]}{Var[X]Var[\Theta]})Var[\Theta]$.
  - *Costruzione di LMS lineare con osservazioni multiple*: $\hat{\Theta}_{LIN} = a_1X_1 + a_2X_2 + ... + a_nX_n + b$.
    - Se: $X_i = \Theta + W_i \to \Theta \perp W_1 \perp W_2 \perp ... \perp W_n$;
    - con: $E[\Theta] = \mu$, $Var[\Theta] = \sigma_\Theta^2$, $E[W_i] = 0$, $Var[W_i] = \sigma_i^2$;
    - allora: $\hat{\Theta}_{LIN} = \frac{\frac{\mu}{\sigma_\Theta^2} + \sum_{i=1}^n \frac{X_i}{\sigma_i^2}}{\sum_{i=0}^n \frac{1}{\sigma_i^2}}$.
    - N.B.: Se $X_i \sim \mathcal{N}$, allora: $\hat{\Theta}_{LIN} = \hat{\Theta}_{LMS} \sim \mathcal{N}$.
- **Stima di $\Theta \sim \mathbf{N}(X_0, \sigma_0^2)$ con $[X_i|\Theta = \theta] \sim \mathbf{N}(\theta, \sigma_i^2)$**:
  - Quando tutte le V.A sono Gaussiane ($X$ e $\Theta$): $\hat{\Theta}_{MAP} = \hat{\Theta}_{LMS} = \hat{\Theta}_{LIN} = m$
  - $m = \frac{\sum_{i=0}^n \frac{X_i}{\sigma_i^2}}{\sum_{i=0}^n \frac{1}{\sigma_i^2}}$

## Simulazione numerica di Esperimenti Aleatori

- **Introduzione**
  - Dato un evento $A$ e un esperimento aleatorio, si vuole stimare la probabilità di $P(A)$. Ripetendo $n$ volte l'esperimento (Montecarlo), la media campionaria ottenuta presenta il seguente errore relativo di stima MSE: $\frac{\sqrt{Var[M_n]}}{P(A)} = \frac{1}{\sqrt{n}}\sqrt{\frac{1-P(A)}{P(A)}} \leq^! \varepsilon$ e dunque $n \geq^! \frac{1}{\varepsilon}\frac{1-P(A)}{P(A)}$.
  - Dato un generatore di bit (pseudo-)casuali, la loro rappresetazione decimale, su una base opportuna sarà distribuita uniformemente tra $U \sim \mathcal{U}[0,1]$.
  - Data $U \sim \mathcal{U}[0,1]$ voglio campionare da una $X \sim f_X$ nota a priori:
  - *Metodo della cumulata inversa*: calcolo $F_X(x)$ e $F_X^{-1}$ ponendo $F_X^{-1} = u$ e ricavando $x \to X = F_X^{-1}(U) \sim f_X(x)$.
  - *App. Campionamento da esponenziale*: $X \sim \mathcal{E}(\lambda) = -\frac{1}{\lambda}\ln U$.
- **Acceptance-Rejection**
  - *Versione "classica"*: $X$ v.a. di legge nota a priori; scelgo un valore $m$ t.c. $m \geq \max(f_X(x))$.
    1: Genero $U \sim \mathcal{U}[0, \max x(x)]$ (ascissa).
    2: Genero $U' \sim \mathcal{U}[0,1]$ in modo indipendente da $U$.
    3: **if** $mU' \leq f_X(U)$ **then**
    4: Accetto e pongo $X = U$.
    5: **else**
    6: Torno a passo 1.
    7: **end if**
  - *Versione "generalizzata"*: Voglio trovare un valore di $m \in \mathbb{N}^+$ t.c. $mf_Y(x) \geq f_X(x)\ \forall x \in \mathbb{N}$.
    1: Genero $Y \sim f_Y$.
    2: Genero $U' \sim \mathcal{U}[0,1]$. $(U' \perp Y)$.
    3: **if** $mf_Y(y)U' \leq f_X(y)$ **then**
    4: Accetto e pongo $X = Y$.
    5: **else**
    6: Torno a passo 1.
    7: **end if**
    - Se $X \sim \mathcal{N}(0,1)$ mi converrebbe che $Y \sim \mathcal{E}(1)$ e $m = 2\sqrt{\frac{e}{2\pi}}$.
      Condizione di acceptance: $U' \leq \frac{f_{|Z|}(Y)}{mf_Y(Y)} = \exp(-\frac{(Y-1)^2}{2})$, ricorda di generare $S$: funzione segno.
    - Per $X \sim \mathcal{N}(\mu, \sigma^2)$ campiono da $Z \sim \mathcal{N}(0,1)$ e pongo $X = \sigma Z + \mu$.
  - *Efficienza degli algoritmi*: $P\left(U \leq \frac{f_X(y)}{mf_Y(y)}\right) = \frac{1}{m}$.
    Oss: $N$: numero prove fino a generare un $x$ valido, $N \sim \mathcal{G}(t)$ con $t$: prob. di acceptance per la singola prova $t = 1/m$.
- **Stima Monte Carlo**
  - Sia $g(X)$ una qualsiasi statistica di $X$, voglio stimare $E[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx$.
  - Per fare ciò uso la media campionaria $\hat{G_n} = \frac{1}{n}\sum_{i=1}^n g(x_i)$ che grazie alla L.L.N.: $\hat{G_n} \to^P E[\hat{G_n}] = E[g(x)]$.
  - Algoritmo:
    1: Genero $X_i \sim f_X$ con $i = 1, ..., n$ in maniera indipendente.
    2: Calcolo $\hat{G_n} = \frac{1}{n}\sum_{i=1}^n g(x_i)$.
  - Errore relativo di stima: visto nell'introduzione, noto che per eventi rari $n$ schizza a valori enormi.
  - Calcolo di integrali: possiamo usare una stima Monte Carlo per il calcolo degli integrali complessi: $I = \int_D g(t)dt = \int_D \frac{g(t)}{f_X(t)}f_X(t)dt = E[\frac{g(X)}{f_X(X)}]$.
    Stimo dunque $\hat{I} = \frac{1}{n}\sum_{i=1}^n \frac{g(X_i)}{f_X(X_i)}$ che per la L.L.N. tenderà in prob. a I.
- **Importance Sampling**
  - Voglio stimare $P(A)$, mediante Montecarlo, ma l'evento $A$ è troppo raro e il numero di sample $n$ è ingestibile; oppure, voglio mantenere $n$ invariato ma cercare di abbassare l'errore relativo di stima. Posso modificare l'esperimento aleatorio in modo da aumentare $P(A)$ e, in seguito, traslerò poi le informazioni ottenute al caso originale.
  - $P_X(A) = E[\mathbb{1}(X \in A)] = E\left[\mathbb{1}(Y \in A)\frac{f_X(y)}{f_Y(y)}\right]$.
  - Algoritmo:
    1: Genero $Y_i \sim f_Y$ i.i.d.
    2: Calcolo $\hat{P_X(A)} = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(Y_i \in A)\frac{f_X(y_i)}{f_Y(y_i)}$
  - Varianza della stima: $Var[\hat{P_X(A)}] = \frac{1}{n}(E[\mathbb{1}(X \in A)\frac{f_X(x)}{f_Y(x)}] - (P_X(A))^2)$.
    Noto che devo scegliere $f_Y(x)$ molto simile a $\mathbb{1}(X \in A)f_X(x)$ per abbassare la varianza della stima.

## Teoria dell'Informazione

- *Informazione di un evento*: $i(A) = \log_2 \frac{1}{P(A)}$.
- *Informazione di eventi indipendenti*: $A \perp B$, $i(A \cap B) = i(A) + i(B)$.
- *Entropia*: $H(X) = E[i(X)] = E\left[\log_2 \frac{1}{p_X(x)}\right] = E[-\log_2 p_X(x)] = -\sum_{j=i}^n p_X(x_j)\log_2 p_X(x_j) \leq \log_2 n$. $H(X) \geq 0\ \forall X$.
- *Diseguaglianza di Jensen*: $\sum_{j=1}^m \lambda_j \log_2 x_j \leq \log_2 \sum_{j=1}^m \lambda_j x_j$; $\lambda_1, \lambda_2, \ldots, \lambda_m : \sum_{j=1}^m \lambda_j = 1$, $\lambda_j \geq 0$. (Valida in generale per funzioni concave).
- *Diseguaglianza di Kraft-McMillan*: Posso trovare un codice prefix-free composto da m codewords con lunghezze $l_j$ $\sum_{j=1}^m 2^{-l_j} \leq 1$.
- *Codifica di sorgente lossless*: $X$ v.a. con $m$ risultati, per ogni codice prefix-free che usa sequenza $l_j$ bit a risultato $X = x_j \to E[L] = \sum_{j=1}^m l_j p_X(x_j) \geq H(X)$.
  Dunque la lunghezza media minima raggiungibile è l'entropia (raggiungibile ad esempio con $n$ prove in modo da poter codificare più lanci insieme, non con 1 sola).
- Con una prova sola la minima lunghezza per rappresentare uno dei risultati è $H(X)$ approssimata per eccesso al primo intero.
- Strategia per costruire un codice prefix-free: assegnare una lunghezza $l_j$ in base a $\lceil i(x_j) \rceil$.