Probabilità totale in una PMF	$p_X\left(x\right)=\sum_{m=1}^{\vert M\vert}\mathbb{P}(E_m)p_{X\mid E_m}(x),\,\bigcup_{i=1}^{m}E_i=\Omega\wedge(E_i\cap E_j=\emptyset,\forall i\neq j\in\{1,..,m\})$
Probabilità Condizionata	$p_X(x)p_{Y\mid X}(y\mid x)=p_{X,Y}(x,y)=p_Y(y)p_{X\mid Y}(x\mid y)$ cond *fisso* $\Rightarrow p_{X\mid Y}(x\mid y)$ è *legge di prob*
Probabilità var indipendenti	$p_{Y\mid X}(y\mid x)=p_Y(y)\qquad p_{X\mid Y}(x\mid y)=p_X(x)$
Regola della Catena	$p_{X,Y,Z}(x,y,z)=p_X(x)p_{Y\mid X}(y\mid x)p_{Z\mid X,Y}(z\mid x,y)$
Linearità della Media	$\mathbb{E}[aX+bY]=a\mathbb{E}[X]+b\mathbb{E}[Y]$
Media Condizionata 1 var	$\mathbb{E}\left[X\right]=\sum_{m=1}^{\vert M\vert}\mathbb{P}(E_m)\sum_{x\in\mathcal{X}}xp_{X\mid E_m}(x\mid E_m)=\sum_{m=1}^{\vert M\vert}\mathbb{P}(E_m)\mathbb{E}\left[X\mid E_m\right]$
Media Condizionata 2 var	$\mathbb{E}[g(X,Y)]=\sum_{y\in\mathcal{Y}}p_Y(y)\sum_{x\in\mathcal{X}}g(x,y)p_{X\mid Y}(X\mid Y)=\sum_{y\in\mathcal{Y}}\mathbb{E}[g(X,Y)\mid Y]p_Y(y)=$ $\mathbb{E}[\mathbb{E}[g(X,Y)\mid Y]]$
Media di funzioni	$\mathbb{E}\left[g(x)\right]=\sum_{x\in\mathcal{X}}g(x)p_X(x)$
Quantità importanti 1 var	$X_{rms}^2=\mathbb{E}\left[X^2\right]=\sum_{x\in\mathcal{X}}x^2p_X(x)\qquad\sigma_X^2=\mathbb{E}\left[\left(X-\mu_X\right)^2\right]=X_{rms}^2-\mu_X^2$
Quantità importanti 2 var	$R_{X,Y}=\mathbb{E}[XY]\qquad COV[X,Y]=\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]=R_{X,Y}-\mu_X\mu_Y$ $COV(X,Y)=\rho_{X,Y}\sigma_X\sigma_Y,\rho\in[-1,1]\qquad p_{X,Y}(x,y)=p_X(x)p_Y(y)\Rightarrow COV(X,Y)=0$
Combinazione lineare 2 var	$Z=aX+bY+c,\,\sigma_Z^2=a^2\sigma_X^2+b^2\sigma_Y^2+2abCOV(X,Y)$
Disuguaglianza di Chebyshev	$X>0,\,\mathbb{P}\left\{\lvert X-\mu_X\rvert\leq k\sigma_X\right\}\geq1-\frac{1}{k^2}$
PMF Congiunta $\Rightarrow$ marginali	dato $p_{X,Y}(x,y),\;p_X(x)=\sum_{y\in\mathcal{Y}}p_{X,Y}(x,y)$
Funzione gen Momenti $\mathbf{M_X}$(s)	$M_X(s)=\mathbb{E}[e^{sX}]\qquad M_X(0)=1\qquad\frac{d^r}{ds^r}M_X(s)\vert_{s=0}=\mathbb{E}[X^r]$ $Z=aX+bY,\,M_Z(s)=M_X^a(s)M_Y^b(s)$
Proprietà prob condizionata	dati X,Y staticamente indipendenti e $Z=X+Y$, $p_{Z\mid X}\left(z\mid x\right)=p_Y\left(z-x\right)$

Nome	**PMF**	$\mu\quad\boldsymbol{\sigma^2}$	**Proprietà**
Uniforme discreta $\quad\mathcal{X}=\{1,\ldots,N\}$	$p(k)=\frac{1}{N}$	$\mu=\frac{N+1}{2}\quad\sigma^2=\frac{N^2-1}{12}$	
Bernoulli $\quad\mathcal{X}=\{0,1\}$	$p(k)=\begin{cases}1-p, & k=0\\ p, & k=1\end{cases}$	$\mu=p\quad\sigma^2=p(1-p)$	
Binomiale $\quad X\sim\mathcal{B}(n,p)$ $\mathcal{X}=\{0,1,\ldots,n\}$	$p(k)=\binom{n}{k}(1-p)^{n-k}p^k$	$\mu=np\quad\sigma^2=$ $np(1-p)$	
Geometrica $\quad X\sim\mathcal{G}(p)\quad\mathcal{X}=\mathbb{N}$	$p(k)=(1-p)^{k-1}p$	$\mu=\frac{1}{p}\quad\sigma^2=\frac{1-p}{p^2}$	
Poisson $\quad X\sim\mathcal{P}(\lambda)\quad\mathcal{X}=\mathbb{N}_0$	$p(k)=\frac{\lambda^ke^{-\lambda}}{k!}\quad\lambda>0)$	$\mu=\lambda\quad\sigma^2=\lambda$	$X_1\sim P(\lambda_1),X_2\sim P(\lambda_2)$ $Y=aX_1+bX_2\sim P(a\lambda_1+b\lambda_2)$

Nome	**PDF**	$\mu\quad\boldsymbol{\sigma^2}$	**CDF**
Uniforme continua $\quad X\sim\mathcal{U}(a,b)$ $\mathcal{X}=[a,b]$	$(x)=\begin{cases}\frac{1}{b-a}, & x\in[a,b]\\ 0, & \text{else}\end{cases}$	$\mu=\frac{a+b}{2}\quad\sigma^2=$ $\frac{(b-a)^2}{12}$	$F(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\leq x\leq b\\ 1, & x\geq b\end{cases}$
Esponenziale $\quad X\sim\mathcal{E}(\lambda)$ $\mathcal{X}=\mathbb{R}_0^+$	$f(x)=\lambda e^{-\lambda x}u(x)\quad\lambda>0$	$\mu=\frac{1}{\lambda}\quad\sigma^2=\frac{1}{\lambda^2}$	$F(x)=(1-e^{-\lambda x})u(x)$
Laplaciana $\quad X\sim\mathcal{L}(\lambda)$ $\mathcal{X}=\mathbb{R}$	$f(x)=\frac{\lambda}{2}e^{-\lambda\lvert x\rvert}\quad\lambda>0$	$\mu=0\quad\sigma^2=\frac{2}{\lambda^2}$	$F(x)=\begin{cases}\frac{1}{2}e^{\lambda x}, & x\leq0\\ 1-\frac{1}{2}e^{-\lambda x}, & x\geq0\end{cases}$
Cauchy $\quad X\sim\mathcal{C}(a,b)\quad\mathcal{X}=\mathbb{R}$	$f(x)=\frac{1}{\pi b}\frac{1}{1+\left(\frac{x-a}{b}\right)^2}$	$\mu=$ non definita $\sigma^2=$ non definita $Sym_{[-H,H]}=a$	$F(x)=\frac{1}{2}+\frac{1}{\pi}\arctan\left(\frac{x-a}{b}\right)$
Rayleigh $\quad X\sim\mathcal{R}\left(\beta\right)$ $\mathcal{X}=\mathbb{R}_0^+$	$f\left(x\right)=\frac{x}{\beta^2}e^{-\frac{x^2}{2\beta^2}}u\left(x\right)\quad\beta>0$	$\mu=\beta\sqrt{\frac{\pi}{2}}$ $\sigma^2=\frac{4-\pi}{2}\beta^2$	$F\left(x\right)=1-e^{-\frac{x^2}{2\beta^2}}$

Nome	**PDF**	$\mu\quad\boldsymbol{\sigma^2}$	**CDF**	**Proprietà**
Gaussiana $X_0\sim\mathcal{N}(0,1)$ $X\sim\mathcal{N}(\mu_X,\sigma_X^2)$	$f_X(x)=\frac{1}{\sqrt{2\pi\sigma_X^2}}e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}$	$\mu=\mu$ $\sigma^2=\sigma^2$	$F_X(x)=1-Q(x)$ $Q(x)=\mathbb{P}(X_0\geq x)$	$X\sim\mathcal{N}(\mu_X,\sigma_X^2)\Rightarrow X=\sigma_XX_0+\mu_X$ $X\sim\mathcal{N}(\mu_X,\sigma_X^2)\Rightarrow\mathbb{P}(X\geq x)=Q\left(\frac{x-\mu_x}{\sigma_x}\right)$ $Q(-x)=1-Q(x)$ $\rho_{1,2}=0\leftrightarrow f_{X_1,X_2}(x_1,x_2)=f_{X_1}(x_1)f_{X_2}(x_2)$ marginali $\iff$ congiunta

$Y=g(X),g$ **continua e iniettiva (invertibile)**	**PDF**	**CDF**
crescente	$f_Y(y)=\frac{f_x\left(g^{-1}(y)\right)}{\lvert g'\left(g^{-1}(y)\right)\rvert}$	$F_Y(y)=F_x(g^{-1}(y))$
descrescente	$f_Y(y)=\frac{f_x\left(g^{-1}(y)\right)}{\lvert g'\left(g^{-1}(y)\right)\rvert}$	$F_Y(y)=1-F_x(g^{-1}(y))$

Matrice di covarianza $\mathbf{K_x}$: $\overline{\overline{\mathbf{K}}}_\mathbf{x} := E\left[\left(\overline{X}-\overline{\mu}_x\right)\left(\overline{X}-\overline{\mu}_x\right)^T\right] = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{1,2} \\ \sigma_1\sigma_2\rho_{1,2} & \sigma_2^2 \end{pmatrix}$

Proprietà della Matrice di covarianza $\mathbf{K_x}$

$|\mathbf{K_x}| = \sigma_1^2\sigma_2^2\left(1-\rho_{1,2}^2\right) \geq 0$

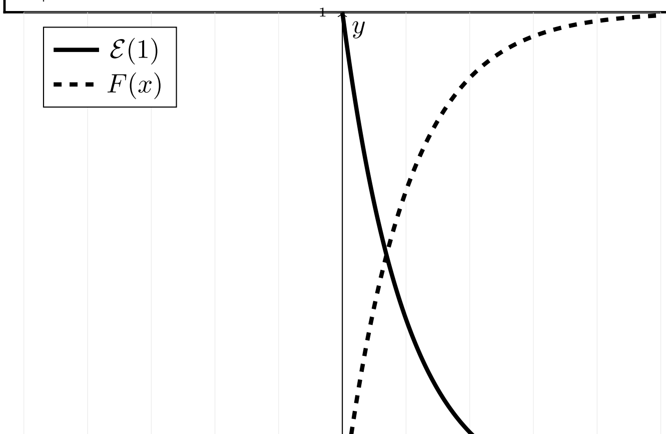
$\rho_{1,2} \neq \pm 1 \implies \mathbf{K_x^{-1}} = \frac{1}{\sigma_1^2\sigma_2^2\left(1-\rho_{1,2}^2\right)}\begin{pmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho_{1,2} \\ -\sigma_1\sigma_2\rho_{1,2} & \sigma_1^2 \end{pmatrix}$

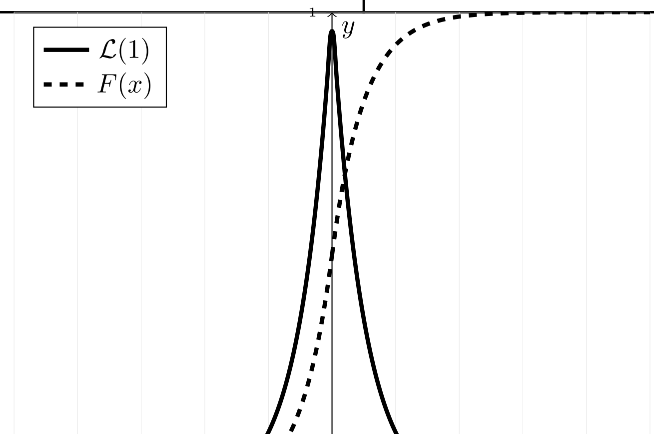
$z = [z_1, z_2]^T \in \mathbb{R}^{2\times 1} \implies \overline{z}^T\overline{\overline{K_x^{-1}}}\,\overline{z} \geq 0$

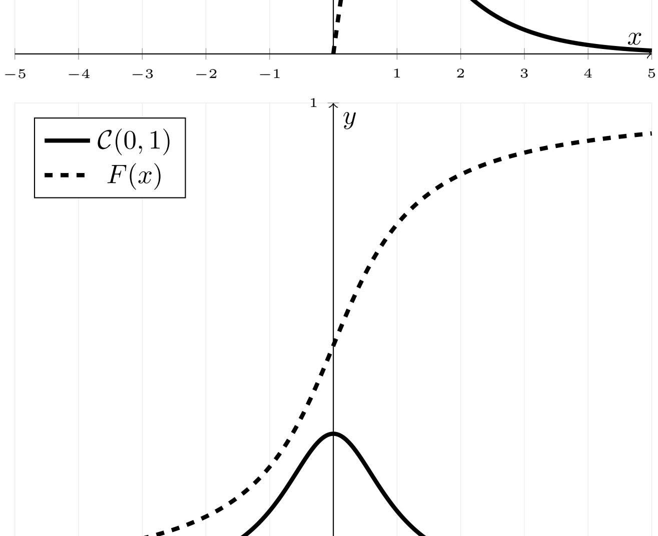
$f_X(\mathbf{x}) = \frac{1}{2\pi|\mathbf{K}_X|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_X)^T\mathbf{K}_X^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)\right] = f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho_{1,2}^2)}}\exp\left[-\frac{\sigma_2^2(x_1-\mu_1)^2+\sigma_1^2(x_2-\mu_2)^2-2\rho_{1,2}(x_1-\mu_1)(x_2-\mu_2)}{2\sigma_1^2\sigma_2^2(1-\rho_{1,2}^2)}\right]$

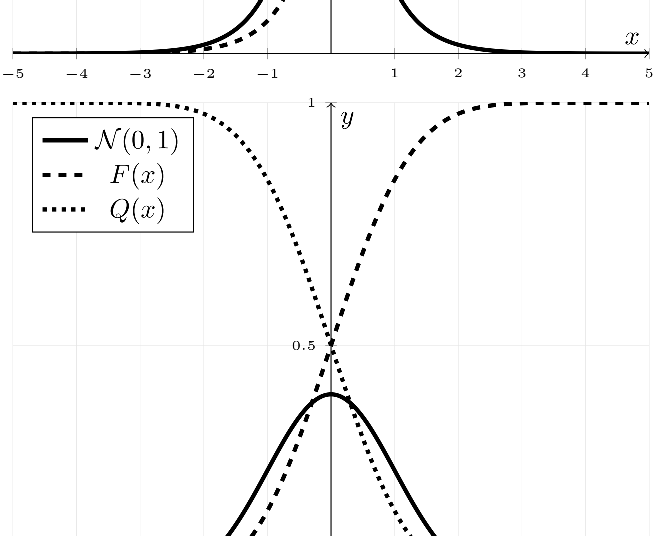
Derivata del Prodotto	$\frac{d}{dx}f(x)\cdot g(x) = f'(x)\cdot g(x) + f(x)\cdot g'(x)$
Derivata del Rapporto	$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)\cdot g(x) - f(x)\cdot g'(x)}{[g(x)^2]}$
Regola della catena	$\frac{d}{dx}f[g(x)] = f'[g(x)]\cdot g'(x)$
Integrazione generale	$\int f\left[g\left(x\right)\right]\cdot g'\left(x\right)dx = F\left[g\left(x\right)\right] + c$
Integrazione per parti	$\int f\left(x\right)\cdot g\left(x\right)dx = F\left(x\right)g\left(x\right) - \int F\left(x\right)\cdot g'\left(x\right)dx$
Funzione Gamma Γ	$\int_0^{+\infty} t^{x-1}e^{-at}dt = a^{-x}\Gamma\left(x\right)$ $\quad\Gamma\left(x\right) := \int_0^{+\infty} t^{x-1}e^{-t}dt$ $\quad\Gamma\left(n\right) = (n-1)!, n\in\mathbb{N}$
Sommatoria primi n numeri	$\sum_{i=1}^n i = \frac{n(n+1)}{2}$ $\quad\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$ $\quad\sum_{i=1}^n i^3 = \left(\frac{n(n+1)}{2}\right)^2$
Serie geometrica	$\sum_{i\in\mathbb{N}_0} p^i = \frac{1}{1-p}, p<1$ $\quad\sum_{i=0}^{n-1} p^i = \frac{1-p^n}{1-p}$ $0\leq p<1, \sum_{n=1}^{+\infty} np^{n-1} = \frac{1}{(1-p)^2}\sum_{n=1}^{\infty} n^2q^{n-1} = \frac{1}{(1-q)^2} + \frac{2q}{(1-q)^3}$

Funzione	**Derivata**	**Integrale**
x^n	nx^{n-1}	$\frac{x^{n+1}}{n+1} + c, n\neq -1$
$\frac{1}{x}$	$-x^{-2}$	$\ln(\lvert x\rvert) + c$
e^x	e^x	$e^x + c$
a^x	$a^x\ln(a)$	$a^x\log_a(e) + c$
$\cos(x)$	$-\sin(x)$	$\sin(x) + c$
$\sin(x)$	$\cos(x)$	$-\cos(x) + c$
$\frac{1}{1+x^2}$	$-\frac{2x}{(1-x^2)^2}$	$\arctan(x) + c$









Definizione informazione	$i(A) = \log\left(\frac{1}{\mathbb{P}(A)}\right) = -\log(\mathbb{P}(A))$
Definizione entropia	$H(X) = \mathbb{E}[i(X)] = \sum_{x\in\mathcal{X}} p_X(x)\log\left(\frac{1}{p_X(x)}\right)$
Entropia Bernoulli	$H(p) = H_2(p) = H(p, 1-p) = p\log\left(\frac{1}{p}\right) + (1-p)\log\left(\frac{1}{1-p}\right) =$ $-p\log(p) - (1-p)\log(1-p)$ $\quad \frac{d}{dp}H(p) = \log\left(\frac{1-p}{p}\right)$
Entropia uniforme discreta	$H(X) = \log(\lvert\mathcal{X}\rvert), \mathcal{X}$ uniforme discreta
Disuguaglianza di Jensen	$g(x)$ differenziabile nei punti dell'alfabeto $\mathcal{X}$ e convessa $\Rightarrow \mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$, $g(x)$ differenziabile nei punti dell'alfabeto $\mathcal{X}$ e concava $\Rightarrow \mathbb{E}[g(X)] \leq g(\mathbb{E}[X])$
Limite superiore Entropia	$H(X) \leq \log(\lvert\mathcal{X}\rvert)$
Divergenza KL (Kullback-Leibler)	$p(x), q(x)$ pmf su $\mathcal{X}$, $D(p\vert\vert q) = \sum_{x\in\mathcal{X}} p(x)\log\left(\frac{p(x)}{q(x)}\right) = \mathbb{E}_{p(x)}\left[\log\left(\frac{p(x)}{q(x)}\right)\right]$, $D(p\vert\vert q) \geq 0 \qquad \forall p, q, D(p\vert\vert q) = 0 \iff p = q$
Disuguaglianza di Kraft	$\sum_{x\in\mathcal{X}} 2^{-l(x)} \leq 1 \iff$ il codice è a prefisso, $l(x)$ lunghezza codice associato a x
Lunghezza di Shannon	$l(x) = \left\lceil \log\left(\frac{1}{p_X(x_i)}\right)\right\rceil$, esiste almeno un codice istantaneo con le lunghezze di Shannon. Inoltre $H(X) \leq L = \mathbb{E}_X[l(X)]$
Entropia coppie variabili	$H(X, Y) = H(X) + H(Y\vert X) = H(Y) + H(X\vert Y)$, $0 \leq H(X\vert Y) \leq H(X), \forall Y$, con $H(X\vert Y) = H(X) \iff X, Y$ indipendenti e $H(X\vert Y) = 0 \iff X = g(Y)$
Probabilità totale Entropia	$H(X\vert Y) = \sum_{y\in\mathcal{Y}} \mathbb{P}(Y = y) H(X\vert Y = y)$
Mutua informazione	$I(X;Y) = H(X) - H(X\vert Y) = H(Y) - H(Y\vert X) = \mathbb{E}\left[\frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)}\right] =$ $D(p_{X,Y}(x, y)\vert\vert p_X(x)p_y(y))$, $\qquad I(X;Y) = 0 \iff p_{X,Y}(x, y) = p_X(x)p_Y(y)$
Entropia vettore aleatorio	$H(X^n) = H(X_1, \ldots, X_n) = H(X_1) + H(X_2\vert X_1) + \ldots + H(X_n\vert X_{n-1}, X_{n-2}, \ldots, X_1)$
Tasso entropico	$H_\infty(X) = \lim\limits_{n\to\infty} \frac{H(\overline{\mathbf{X}}^n)}{n} = \lim\limits_{n\to\infty} H(X_n\vert X_{n-1}, X_{n-2}, \ldots, X_1)$
Sorgenti Markoviane	$p_{X(n)\vert X(n-1),\ldots,X(1)}(x_n\vert x_{n-1}, \ldots, x_1) = p_{X(n)\vert X(n-1)}(x_n\vert x_{n-1})$
Matrice di transizione	dato $\mathcal{X} = \{1, \ldots, m\}, \overline{\overline{P}} \in \mathcal{M}_{m\times m}$ tale che $\forall i, j \in \mathcal{X} : p_{i\to j}$ è la probabilità di andare dallo stato i allo stato j
Vettore di probabilità degli stati a tempo n	$\overline{x}(n_o + n) = \left(p_{X(n_0+n)}(1), \ldots, p_{X(n_0+n)}(m)\right) = \left(\overline{\overline{P^T}}\right)^n \overline{x}(n_0)$
Condizione di stazionarità	$\overline{x}$ è distribuzione stazionaria (indipendente da n) $\iff \overline{\overline{P^T}}\overline{x} = \overline{x}$
Distribuzione stazionaria grafo NON orientato	$\overline{x} = \left(\frac{w_1}{2w}, \ldots, \frac{w_m}{2w}\right), w_i =$ somma pesi archi uscenti da S_i e w somma pesi tutti gli archi.
Tasso entropico Markov	$H_\infty(X) = \sum_{i=1}^m \overline{x}_i \sum_{j=1}^m p_{i\to j}\log\left(\frac{1}{p_{i\to j}}\right) = \sum_{i=1}^m \overline{x}_i H(S_i)$, dove $H(S_i)$ è l'incertezza di uscire dallo stato i

Tipo inferenza	Cose note	Regola decisione/Stimatore
Decisione Bayesiana	$\{H_i\}_{i=1}^{M}$ disgiunte $\forall i \in \{1,\dots,M\} : \mathbb{P}\{H_i\}$ $\forall i \in \{1,\dots,M\} : \mathbb{P}\{X^n = x^n \vert H_i\}$ $\overline{\overline{C}} \in \mathcal{M}_{m\times m} : c_{i,j}$ costo se decido i ma è vera l'ipotesi H_j	$\mathcal{R} = \sum_{i=1}^{M}\sum_{j=1}^{M} C_{i,j}\mathbb{P}\{D(X^n) = i, H = H_j\}$ $D(x^n) = i \iff \mathbb{P}\{X^n = x^n, H_i\} > \mathbb{P}\{X^n = x^n, H_j\} \forall j \neq i$ $L(x^n) = \frac{\mathbb{P}\{X^n=x^n \vert H_1\}}{\mathbb{P}\{X^n=x^n \vert H_2\}} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{\mathbb{P}\{H_2\}}{\mathbb{P}\{H_1\}}$ con 2 ipotesi
Decisione non Bayesiana (Neyman-Pearson)	H_0 stato normale, $H_1 = \overline{H_0}$ stato alterato $\forall i \in \{0,1\} : \mathbb{P}\{X^n = x^n \vert H_i\}$ Errore di tipo I: $P\{D(X^n) = 1 \vert H_0\}$ Potenza del test: $P\{D(X^n) = 1 \vert H_1\}$	$L(x^n) = \frac{\mathbb{P}\{X^n=x^n \vert H_1\}}{\mathbb{P}\{X^n=x^n \vert H_0\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$ $\alpha = \mathbb{P}\{L(x^n) > \eta \vert H_0\}$ $1-\beta = \mathbb{P}\{L(x^n) > \eta \vert H_1\}$
Stima Bayesiana parametro continuo	$f_{\Theta \vert X^n}(\theta \vert x^n) = \frac{f_\Theta(\theta)p_{X^n \vert \Theta}(x^n \vert \theta)}{p_{X^n}(x^n)}$ $p_{X^n}(x^n) = \int f_\Theta(\theta)p_{X^n \vert \Theta}(x^n \vert \theta)d\theta$ $\mathcal{R} = E\left[C\left(\hat{\Theta}(X^n) - \Theta\right)\right] =$ $E_{x^n}\left[E\left[C\left(\hat{\Theta}(X^n) - \Theta\right)\Big\vert X^n\right]\right]$	$\hat{\theta}_{opt}(x^n) = \arg\min \int C(\hat{\theta}(x^n) - \theta)f_{\Theta \vert X^n}(\theta \vert x^n)d\theta$ $C(x) = x^2 \Rightarrow \hat{\theta}_{mmse}(x^n) = \int \theta f_{\Theta \vert X^n}(\theta \vert x^n)d\theta = \mathbb{E}[\Theta \vert X^n = x^n]$ $C(x) = \Pi(\frac{x}{\epsilon}) = \begin{cases} 1, & \vert x \vert < \frac{\epsilon}{2} \\ 0, & else \end{cases} \Rightarrow \hat{\theta}_{map}(x^n) = \arg\max f_{\Theta \vert X^n}(\theta \vert x^n)$
Stima non Bayesiana	$f_{X^n}(x^n;\theta)$	$\hat{\theta}_{ML}(x^n) = \arg\max \log\left(f_{X^n}(x^n;\theta)\right)$

NOTA: APPLICARE IL LOGARITMO NATURALE NON CAMBIA IL VERSO DELLA DISEQUAZIONE, FORTEMENTE CONSIGLIATI IN TUTTI GLI STIMATORI.

Bias/unbias	$\hat{\Theta}$ unbiased/corretto $\iff \mathbb{E}[\hat{\Theta}(X^n) \vert \Theta = \theta] = \theta$
unbias/correto asintotico	$\hat{\Theta}$ unbiased/corretto asintoticamente $\iff \lim_{n\to\infty}(\mathbb{E}[\hat{\Theta}(X^n) \vert \Theta = \theta]) = \theta$
Consistenza in probabilità	$\lim_{n\to\infty} \mathbb{P}\left\{\left\vert\hat{\Theta}(X^n) - \Theta\right\vert > \epsilon\right\} = 0, \forall\epsilon > 0$
Consistenza in media quadratica	$\hat{\Theta}$ consistente in media quadratica $\iff \lim_{n\to\infty}(\overline{e}^2) = 0$, $\overline{e}^2 = \mathbb{E}[(\hat{\Theta}(X^n) - \Theta)^2]$ se $\hat{\Theta}$ è unbiased $\overline{e}^2 = Var(\hat{\Theta}(X^n))$
Informazione di Fisher	$I_n = \mathbb{E}[(\frac{d\log(f_{X^n}(x^n;\theta))}{d\theta})^2] = -\mathbb{E}[\frac{d^2\log(f_{X^n}(x^n;\theta))}{d^2\theta}]$
Limite CR	$Var[\hat{\Theta}(X^n)] \geq \frac{1}{I_n(\theta)}$ con $\hat{\Theta}(X^n)$ efficiente se vale '='

NOTA: Se in uno stimatore vale che:

- $C(\cdot) \geq 0$ e convessa ($\cup$)
- $f_{\Theta \vert X^n}(\theta \vert x^n)$ è simmetrica rispetto a $\mathbb{E}[\Theta \vert X^n = x^n]$
- $f_{X^n}(x^n;\theta)$ è unimodale

Allora tutti gli stimatori che ottimizzano $C(\cdot)$ coincidono con $\mathbb{E}[\Theta \vert X^n]$ e, quindi equivalenti a $\hat{\Theta}_{mmse}(X^n) = \hat{\Theta}_{map}(X^n)$.
NOTA:

- Convessa = $\cup$
- Concava = $\cap$

Nome	**Entropia H**	**Stimatori/Likelihood** dato X^n con n realizzazioni (x_i) iid	**Inform. Fisher $\mathbf{I_n}$**
Uniforme discreta $\mathcal{X} = \{1, \ldots, m\}$	$H = \ln(\lvert\mathcal{X}\rvert) = \ln(m)$	Stimare il parametro $m \implies$ $\hat{m} = \max\{x_1, \ldots, x_n\}$	
Uniforme continua $X \sim \mathcal{U}(a, b)$ $\mathcal{X} = [a, b]$	$H = \log(b - a)$	Stimare il parametro a o $b \implies$ $\hat{a} = \min\{x_1, \ldots, x_n\}$ $\hat{b} = \max\{x_1, \ldots, x_n\}$	
Bernoulli $\mathcal{X} = \{0, 1\}$	$H = -\log(1-p) + p\log\left(\frac{1-p}{p}\right)$	Stimare il parametro $p \implies$ $\Lambda(p) = \log(p)\sum_{i=1}^n x_i + \log(1-p)\sum_{i=1}^n (1 - x_i)$	$I_n = \frac{1}{p(1-p)}$
Binomiale $X \sim \mathcal{B}(n, p)$ $\mathcal{X} = \{0, 1, \ldots, n\}$		Stimare il parametro $p \implies \Lambda(p) = \sum_{i=1}^m \log\binom{n}{x_i} + \log(p)\sum_{i=1}^m x_i + \log(1-p)\sum_{i=1}^m (n - x_i)$	$I_n = \frac{n}{p(1-p)}$
Geometrica $X \sim \mathcal{G}(p)$ $\mathcal{X} = \mathbb{N}$	$H = -\log(p) - \frac{1-p}{p}\log(1-p)$	Stimare il parametro $p \implies$ $\Lambda(p) = n\log(p) + \log(1-p)\sum_{i=1}^n (x_i - 1)$	$I_n = \frac{1}{p^2(1-p)}$
Poisson $X \sim \mathcal{P}(\lambda)$ $\mathcal{X} = \mathbb{N}_0$		Stimare il parametro $\lambda \implies$ $\Lambda(\lambda) = \ln(\lambda)\sum_{i=1}^n x_i - n\lambda - \sum_{i=1}^n \ln(x_i!)$	$I_n = \frac{1}{\lambda}$
Esponenziale $X \sim \mathcal{E}(\lambda)$ $\mathcal{X} = \mathbb{R}_0^+$	$H = 1 - \ln(\lambda)$	Stimare il parametro $\lambda \implies$ $\Lambda(\lambda) = n\ln(\lambda) - \lambda\sum_{i=1}^n x_i$	$I_n = \frac{1}{\lambda^2}$
Laplaciana $X \sim \mathcal{L}(\lambda)$ $\mathcal{X} = \mathbb{R}$	$H = 1 + \ln\left(\frac{2}{\lambda}\right)$	Stimare il parametro $\lambda \implies$ $\Lambda(\lambda) = n\ln(\lambda) - n\ln(2) - \lambda\sum_{i=1}^n \lvert x_i\rvert$	$I_n = \frac{1}{\lambda^2}$
Cauchy $X \sim \mathcal{C}(a, b)$ $\mathcal{X} = \mathbb{R}$	$H = \log(4\pi b)$	Stimare i parametri a e $b \implies$ $\Lambda(a; b) = -n\log(b\pi) - \sum_{i=1}^n \log\left(1 + \left(\frac{x_i - a}{b}\right)^2\right)$	$I_n = \frac{1}{2b^2}$
Rayleigh $X \sim \mathcal{R}(\beta)$ $\mathcal{X} = \mathbb{R}_0^+$	$H = 1 + \ln\left(\frac{\beta}{\sqrt{2}}\right) + \frac{\gamma}{2}$ $\gamma \approx 0.5772\ldots$	Stimare il parametro $\beta \implies$ $\Lambda(\beta) = -2n\ln(\beta) + \sum_{i=1}^n \ln(x_i) - \frac{1}{2\beta^2}\sum_{i=1}^n x_i^2$	$I_n = \frac{2}{\beta^2}$
Gaussiana $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$	$H = \frac{1}{2}\left(1 + \ln\left(2\sigma_X^2\pi\right)\right)$	Stimare i parametri μ_X e $\sigma_X^2 \implies$ $\Lambda(\mu_X; \sigma_X^2) = -\frac{n}{2}\ln\left(2\pi\sigma_X^2\right) - \frac{1}{2\sigma_X^2}\sum_{i=1}^n (x_i - \mu_X)^2$	$I_n^{\mu_X} = \frac{1}{\sigma_X^2}$ $I_n^{\sigma_X^2} = \frac{1}{2\sigma_X^2}$