DIMOSTRARE IL PRINCIPIO DI INCLUSIONE-ESCLUSIONE

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

DIM:

$$P(A \cup B) = P(A \cup (B \setminus A))$$
$$= P(A \cup (B \cap A^c))$$
$$= P(A) + P(B \cap A^c)$$
$$= P(A) + P(B) - P(A \cap B)$$

SAPENDO CHE

$$P(B) = P(B \cap A) + P(B \cap A^c)$$
$$P(B \cap A^c) = P(B) - P(B \cap A)$$

DIMOSTRARE LA FORMULA DELLE PROBABILITÀ TOTALI

DATA $A_1 ... A_m$ PARTIZIONE DI Ω

$$P(B) = \sum_{i=1}^{m} P(B|A_i) \cdot P(A_i)$$

PER DEFINIZIONE:

$$P(B|A_i) = \frac{P(B \cap A_i)}{P(A_i)}$$

SAPENDO CHE

$$B = \bigcup (B \cap A_i)$$
$$P(B) = \sum_{i=1}^{m} P(B \cap A_i)$$

$$\Rightarrow P(B \cap A_i) = P(B|A_i) \cdot P(A_i)$$
$$\Rightarrow P(B) = \sum_{i=1}^{m} P(B \cap A_i) = \sum_{i=1}^{m} P(B|A_i) \cdot P(A_i)$$

DIMOSTRARE $VAR(X) = E[X^2] - E[X]^2$

$$VAR(X) = \frac{\sum_{i=1}^{m} (x_i - E[X])^2}{m} = \frac{\sum_{i=1}^{m} x_i^2 - 2x_i E[X] + E[X]^2}{m}$$

$$= \sum_{i=1}^{m} \frac{x_i^2}{m} - \frac{\sum_{i=1}^{m} 2x_i E[X]}{m} + \frac{\sum_{i=1}^{m} E[X]^2}{m} = E[X^2] - 2E[X] \cdot E[X] + E[X]^2$$

$$= E[X^2] - E[X]^2$$

CONFIDENZA SIA $0<c<1$

$$P(-z_c < z_0 < z_c) = c$$

X_i $i=1,\dots,m$ = SUCCESSIONE DI V.A. CON LA STESSA DENSITA DI MEDIA μ e VARIANZA σ^2

SIA $\overline{X}_m = \dfrac{X_1+\dots+X_m}{m}$, PER IL TEOREMA DEL LIMITE CENTRALE:

$$E[\overline{X}_m] = \mu \quad VAR(\overline{X}_m) = \frac{\sigma^2}{m}$$

$$\overline{X}_m \sim N\left(\mu, \frac{\sigma^2}{m}\right)$$

$$P\left(-z_c < \frac{X_m - \mu}{\frac{\sigma}{\sqrt{m}}} < z_c\right) = c$$

$$= P\left(-z_c \cdot \frac{\sigma}{\sqrt{m}} < X_m - \mu < z_c \cdot \frac{\sigma}{\sqrt{m}}\right) = c$$

$$= P\left(-X_m - z_c \frac{\sigma^2}{\sqrt{m}} < -\mu < -X_m + z_c \cdot \frac{\sigma}{\sqrt{m}}\right) = c$$

$$= P\left(X_m - z_c \cdot \frac{\sigma}{\sqrt{m}} < \mu < X_m + z_c \frac{\sigma}{\sqrt{m}}\right) = c$$

APPROSSIMIAMO σ CON S = SCARTO MEDIO QUADRATICO CAMPIONARIO

$$= \mathrm{Conf}\left(X_m - z_c \cdot \frac{S}{\sqrt{m}} < \mu < X_m + z_c \cdot \frac{S}{\sqrt{m}}\right) = c$$

DIMOSTRARE $Cov(XY) = E[XY] - E[X] \cdot E[Y]$

$$Cov(XY) = \frac{\sum_{i=1}^{m} (x_i - E[X]) \cdot (y_i - E[Y])}{m}$$

$$= \frac{1}{m} \sum_{i=1}^{m} \left(x_i y_i - x_i E[Y] - y_i E[X] + E[X] \cdot E[Y] \right)$$

$$= E[XY] - E[X] \cdot E[Y] - \cancel{E[X] \cdot E[Y]} + \cancel{E[X] \cdot E[Y]}$$

Se $Y = aX + b$ $\quad E[Y] = a\,E[X] + b$

$$\sum_{i=1}^{m} \frac{ax_i + b}{m} = \frac{1}{m} \left(\sum_{i=1}^{m} a x_i + \sum_{i=1}^{m} b \right) = a\,E[X] + b$$

Se $Y = aX + b$ $\quad VAR(Y) = a^2 \cdot VAR(X)$

$$\frac{1}{m} \sum_{i=1}^{m} \left(a x_i + b - E[Y] \right)^2 = \frac{1}{m} \cdot \sum_{i=1}^{m} \left(a x_i + b - (a\,E[X] + b) \right)^2$$

$$= \frac{1}{m} \sum_{i=1}^{m} \left(a x_i + \cancel{b} - a\,E[X] - \cancel{b} \right)^2 = \frac{1}{m} \cdot \sum \left(a \cdot (x_i - E[X]) \right)^2$$

$$= \frac{a^2}{m} \cdot \sum \left(x_i - E[X] \right)^2 = a^2 \cdot VAR(X)$$

DIMOSTRARE LA FORMULA DI Bayes

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

SAPENDO CHE

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}}$$

$$P(A \cap B) = P(A \mid B) \cdot P(B)$$

$$P(A \cap B) = P(B \mid A) \cdot P(A)$$

$$\Rightarrow P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)$$

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

MEDIA E VARIANZA DI UNA VARIABILE DISCRETA UNIFORME 4

$$X \sim U(1, \dots, m)$$

$$P_X(K) = \begin{cases} \frac{1}{m} & \text{se } K \in (1, \dots, m) \\ 0 & \text{ALTRIMENTI} \end{cases}$$

$$E[X] = \sum_{K=1}^{m} \frac{1}{m} \cdot K = \frac{1}{m} \sum_{K=1}^{m} K = \frac{1}{\cancel{m}} \cdot \frac{\cancel{m} \cdot (m+1)}{2} = \frac{m+1}{2}$$

$$E[X^2] = \sum_{K=1}^{m} \frac{1}{m} K^2 = \frac{1}{m} \cdot \sum_{K=1}^{m} K^2 = \frac{1}{\cancel{m}} \cdot \frac{\cancel{m} \cdot (m+1) \cdot (2m+1)}{6}$$

$$VAR(X) = E[X^2] - E[X]^2 = \frac{(m+1) \cdot (2m+1)}{6} - \frac{(m+1)^2}{4}$$

$$= \frac{2(m+1) \cdot (2m+1) - 3(m+1)^2}{12} = \frac{2 \cdot (2m^2 + m + 2m + 1) - 3 \cdot (m^2 + 2m + 1)}{12}$$

$$= \frac{4m^2 + \cancel{6m} + 2 - 3m^2 - \cancel{6m} - 3}{12} = \frac{m^2 - 1}{12}$$

MEDIA E VARIANZA DI UNA VARIABILE ALEATORIA CONTINUA UNIFORME SU $[c, d]$

$$f(x) = \begin{cases} \frac{1}{d-c} & \text{se } c < x < d \\ 0 & \text{ALTRIMENTI} \end{cases}$$

SAPENDO CHE $a^3 - b^3 = (a-b) \cdot (a^2 + ab + b^2)$

$$E[X] = \int_c^d x \cdot \frac{1}{d-c} dx = \frac{1}{d-c} \cdot \left[\frac{x^2}{2} \right]_c^d = \frac{1}{2(d-c)} \cdot d^2 - c^2 = \frac{(d+c) \cdot \cancel{(d-c)}}{2 \cdot \cancel{(d-c)}}$$

$$= \frac{(d+c)}{2}$$

$$E[X^2] = \int_c^d x^2 \cdot \frac{1}{d-c} dx = \frac{1}{d-c} \left[\frac{x^3}{3} \right]_c^d = \frac{1}{3 \cdot (d-c)} (d^3 - c^3)$$

$$= \frac{\cancel{(d-c)} \cdot (d^2 + cd + c^2)}{3 \cdot \cancel{(d-c)}}$$

$$VAR(X) = E[X^2] - E[X]^2 = \frac{d^2 + cd + c^2}{3} - \frac{(d+c)^2}{4} = \frac{4d^2 + 4cd + 4c^2 - 3d^2 - 3c^2 - 6cd}{12}$$

$$= \frac{d^2 - 2cd + c^2}{12} = \frac{(d-c)^2}{12}$$

DIMOSTRARE MEDIA E VARIANZA DI VARIABILI ESPONENZIALI

$X \sim Exp(\lambda)$

$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{se } t \geq 0 \\ 0 & \text{ALTRIMENTI} \end{cases}$$

$$E[X] = \int_0^{+\infty} x \cdot \lambda \cdot e^{-\lambda x} dx = \lambda \cdot \left(\left[x \frac{e^{-\lambda x}}{-\lambda} \right]_0^{\infty} - \int_0^{\infty} \frac{e^{-\lambda x}}{\lambda} dx \right)$$

$$= \cancel{\lambda} \cdot \left(\left[x \cdot \frac{e^{-\lambda x}}{\cancel{-\lambda}} - \frac{1}{\cancel{\lambda}} \cdot \frac{e^{-\lambda x}}{-\lambda} \right]_0^{\infty} \right) = \frac{1}{\lambda}$$

$$E[X^2] = \int_0^{\infty} x^2 \cdot \lambda \cdot e^{-\lambda x} dx = \cancel{\lambda} \cdot \left(\left[-x^2 \cdot \frac{e^{-\lambda x}}{+\cancel{\lambda}} \right]_0^{\infty} + \int_0^{\infty} 2x \cdot \frac{e^{-\lambda x}}{+\cancel{\lambda}} dx \right)$$

$$= \cancel{\left[-x^2 e^{-\lambda x} \right]_0^{\infty}} + 2 \cdot \left(\left[x \cdot \frac{e^{-\lambda x}}{-\lambda} \right]_0^{\infty} - \int_0^{\infty} \frac{e^{-\lambda x}}{-\lambda} dx \right)$$

$$= 2 \cdot \left(\left[x \cdot \frac{e^{-\lambda x}}{-\lambda} + \frac{1}{\lambda} \cdot \frac{e^{-\lambda x}}{-\lambda} \right]_0^{\infty} \right) = 2 \cdot \left[-\frac{1}{\lambda^2} \cdot e^{-\lambda x} \right]_0^{\infty} = \frac{+2}{\lambda^2}$$

$$VAR(X) = E[X^2] - E[X]^2 = +\frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

DIMOSTRARE CHE LA MEDIA RENDE MINIMA LA VARIANZA

$$\sigma_X^2 = \frac{\sum_{i=1}^{m} (x_i - t)^2}{m} = \frac{1}{m} \cdot (x_1 - t)^2 + \dots + (x_m - t)^2$$

$$f'(t) = \frac{2}{m} \cdot \left((t - x_1) \cdot (t - x_2) \cdot \dots \cdot (t - x_m) \right)$$

$$= \frac{2}{m} \cdot \left(mt - (x_1 + x_2 + \dots + x_m) \right) = 0$$

$$t = \frac{(x_1 + x_2 + \dots + x_m)}{m} = \overline{x_m}$$

RETTA DI REGRESSIONE

$$y = ax + b \qquad \begin{cases} \bar{y} = a\bar{x} + b \\ a = \dfrac{\sigma_{xy}}{\sigma^2_x} \end{cases}$$

SIA $S(a,b)$ LA SOMMA DEI QUADRATI DEGLI ERRORI

$$S(a,b) = \sum_{i=1}^{n} (y_i - ax - b)^2$$

SAPENDO CHE

$$\boxed{(a+b+c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc}$$

AGGIUNGO E TOLGO $a\bar{x}$ e $\bar{y}$

$$= \sum_{i=1}^{n} \Big(\underbrace{(y_i - \bar{y})}_{1} - \underbrace{(ax - a\bar{x})}_{2} + \underbrace{(-b - a\bar{x} + \bar{y})}_{3} \Big)^2$$

$$= \sum_{i=1}^{n} \Big((y_i - \bar{y})^2 + (ax - a\bar{x})^2 + (-b - a\bar{x} + \bar{y})^2 - 2 \cdot (y_i - \bar{y}) \cdot (ax - a\bar{x}) \Big)$$

$$= n \cdot \sigma_y^2 + a^2 n \sigma^2_x - 2na\sigma_{xy} + (-b - a\bar{x} + \bar{y})^2$$

IMPONIAMO $-b - a\bar{x} + \bar{y} = 0 \Rightarrow \bar{y} = a\bar{x} + b$

TROVIAMO IL MINIMO DI $a^2 n \sigma^2_x - 2n\sigma_{xy} a + n \cdot \sigma^2_y$

$$f'(a) = 2n\sigma^2_x a - 2n\sigma_{xy} = 0$$

$$a = \frac{\cancel{2n}\sigma_{xy}}{\cancel{2n}\sigma^2_x} \Rightarrow a = \frac{\sigma_{xy}}{\sigma^2_x}$$

DISUGUAGLIANZA DI CHEBISHEV

$$\mathbb{P}\left(|X-E[X]|>\eta\right) \leq \frac{Var(X)}{\eta^2}$$

$$A=\{|X-E[X]|>\eta\}$$

SIANO:

$$Y=\eta^2\chi_A \qquad \chi_A(\omega)=\begin{cases}1 & \text{se } \omega\in A\\ 0 & \text{ALTRIMENTI}\end{cases}$$

$$Z=\left(X-E[X]\right)^2 \qquad \chi_A \sim B(1,\mathbb{P}(A))$$

Se $\omega \in A$

$$\Rightarrow Y(\omega)=\eta^2 \quad \text{e} \quad Z(\omega)>\eta^2$$

Se $\omega \notin A$

$$\Rightarrow Y(\omega)=0 \quad \text{e} \quad Z(\omega)\geq 0 \quad (\text{PERCHÈ È UN QUADRATO})$$

$$\Rightarrow Y(\omega) \leq Z(\omega)$$

$$E[Y] \leq E[Z] = E[(X-E[X])^2] = Var(X)$$

$$E[Y] = \eta^2\cdot E[\chi_A] = \eta^2\cdot\mathbb{P}(A) \qquad \Rightarrow \eta^2\cdot\mathbb{P}(A) \leq Var(X)$$

$$\Rightarrow \mathbb{P}\left(|X-E[X]|>\eta\right) \leq \frac{Var(X)}{\eta^2}$$

CONVERGENZA IN PROBABILITÀ

CONSIDERIAMO UNA SUCCESSIONE DI VARIABILI ALEATORIE $X_1, X_2, \dots$.

DICIAMO CHE $X_1, X_2, \dots X_n$ CONVERGONO IN PROBABILITÀ A X SE $\forall \eta>0$ FISSATO

$$\lim_{n\to\infty}\mathbb{P}\left(|X_n-X|>\eta\right)=0$$

LEGGE DEI GRANDI NUMERI

SIANO $X_1, \dots, X_m$ V.A. INDIPENDENTI AVENTI LA STESSA DENSITÀ. CON $E[x_i] = \mu \ \forall i$ $Var(x_i) = \sigma^2$

SIA $\bar{X}_m = \frac{X_1 + \dots + X_m}{m}$ N.B. $\bar{X}_m$ È A SUA VOLTA UNA V.A.

ALLORA $\bar{X}_m$ CONVERGE IN PROBABILITÀ A μ

DIMOSTRAZIONE:

$$E[\bar{X}_m] = E\left[\frac{X_1 + \dots + X_m}{m}\right] = \frac{1}{m}\cdot\left(E[X_1] + \dots + E[X_m]\right) = \frac{1}{\cancel{m}}\cdot\cancel{m}\cdot \underset{\mu}{E[X_i]} = \mu$$

$$Var(\bar{X}_m) = Var\left(\frac{X_1 + X_2 + \dots + X_m}{m}\right) = \frac{1}{m^2}\cdot\left(Var(X_1) + \dots Var(X_m)\right) = \frac{1}{m^{\cancel{2}}}\cdot\cancel{m}\cdot\sigma^2 = \frac{\sigma^2}{m}$$

APPLICHIAMO LA DISUGUAGLIANZA DI CHEBISHEV $\left(P(|X - E[X]| > \eta) \leq \frac{Var(x)}{\eta^2}\right)$

$$P(|X_m - \mu| > \eta) \leq \frac{\frac{\sigma^2}{m}}{\eta^2} \xrightarrow[m \to +\infty]{} 0$$ (ALLORA X_m CONVERGE A μ IN PROBABILITÀ)

MEDIA GEOMETRICA O LOGARITMICA $f = \log$ $f^{-1} = e$

$$= f^{-1}\left(\frac{f(x_1) + f(x_2) + \dots + f(x_m)}{m}\right)$$

$$= e^{\left(\frac{\log(x_1) + \dots + \log(x_m)}{m}\right)} = e^{\frac{1}{m}\cdot(\log(x_1) + \log(x_2) + \dots \log(x_m))}$$

$$= e^{\frac{1}{m}\cdot(\log(x_1\cdot x_2\cdot \dots \cdot x_m))} = e^{\log\left((x_1\cdot x_2 \cdot \dots \cdot x_m)^{\frac{1}{m}}\right)}$$

$$= \sqrt[m]{x_1\cdot x_2 \cdot \dots \cdot x_m}$$

FUNZIONE DI RIPARTIZIONE DI UNA VARIABILE GEOMETRICA MODIFICATA $\tilde{G}$

$X \sim \tilde{G}(p)$

$$p_X(k) = \begin{cases} p\cdot(1-p)^{k-1} & \text{se } k=1,\dots \\ 0 & \text{ALTRIMENTI} \end{cases}$$

– SAPENDO CHE –

$$\boxed{\sum_{k=0}^{m} x^k = \frac{x^{m+1}-1}{x-1}}$$

FUNZIONE DI RIPARTIZIONE:

$$P(X\leq x) = \sum_{k=1}^{x} p\cdot(1-p)^{k-1} = p\cdot\sum_{k=1}^{x}(1-p)^{k-1} = p\cdot\sum_{k=0}^{x-1}(1-p)^k$$

$$= p\cdot\frac{(1-p)^x-1}{(\cancel{1}-p)-\cancel{1}} = 1-(1-p)^x$$

MANCANZA DI MEMORIA DELLE VARIABILI GEOMETRICHE

$X \sim \tilde{G}(p)$

$$P(X=t+s \mid X\geq s) = P(X=t)$$

PERCHÈ PIÙ RESTRITTIVA

PER DEFINIZIONE $\quad \| \quad \dfrac{P(X=t+s \cap X\geq s)}{P(X\geq s)} = \dfrac{P(X=t+s)}{1-P(X\leq s)}$

$$= \frac{p\cdot(1-p)^{t+s}}{\cancel{1}-(\cancel{1}-(1-p)^s)} = \frac{p\cdot(1-p)^t\cdot\cancel{(1-p)^s}}{\cancel{(1-p)^s}} = P(X=t)$$

MEDIA DI UNA VARIABILE GEOMETRICA MODIFICATA $E[X]=\frac{1}{p}$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \Rightarrow \sum_{k=1}^{\infty} x^k = \frac{1}{1-x}-1 = \frac{\cancel{1}-\cancel{1}+x}{1-x} = \frac{x}{1-x}$$

$$\text{se } f(x) = \sum_{k=1}^{\infty} x^k \Rightarrow f'(x) = \frac{1\cdot(1-x)-(x)\cdot(-1)}{(1-x)^2} = \frac{1}{(1-x)^2}$$

$$E[X] = \sum_{k=1}^{\infty} k\cdot p\cdot(1-p)^{k-1} = p\cdot\sum_{k=1}^{\infty} k\cdot(1-p)^{k-1} = p\cdot f'(1-p)$$

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

DIMOSTRARE $E[X+Y] = E[X] + E[Y]$

$\phi(X,Y) = X+Y \qquad E[X+Y] = E[\phi(X,Y)]$

$$= \sum_{(x,y)\in\mathbb{R}^2} (x+y) \cdot P_{(X,Y)}(x,y)$$

$$= \sum_{(x,y)\in\mathbb{R}^2} x \cdot P_{(X,Y)}(x,y) + \sum_{(x,y)\in\mathbb{R}^2} y \, P_{(X,Y)}(x,y)$$

$$= \sum_{x\in\mathbb{R}} x \cdot \underbrace{\sum_{y\in\mathbb{R}} P_{X,Y}(\overset{\text{x FISSATO}}{x},y)}_{P_X(x)} + \sum_{y\in\mathbb{R}} y \cdot \underbrace{\sum_{x\in\mathbb{R}} P_{X,Y}(x,\overset{\text{Y FISSATO}}{y})}_{P_Y(y)}$$

$$= E[X] + E[Y]$$

FUNZIONE DI RIPARTIZIONE DI $X \sim Exp(\lambda)$

$$p_X(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{se } t>0 \\ 0 & \text{se } t\leq 0 \end{cases}$$

$$F(x) = \int_0^x \lambda \cdot e^{-\lambda t} dt = \cancel{\lambda} \cdot \left[\frac{e^{-\lambda t}}{-\cancel{\lambda}}\right]_0^x = -e^{-\lambda x} + 1 = 1 - e^{-\lambda x}$$

MANCANZA DI MEMORIA DELLE VARIABILI ESPONENZIALI

$X \sim Exp(\lambda)$ $\qquad P(X>t+s \mid X>t) = P(X>s)$ PIÙ RESTRITTIVO DI $P(X>t)$

PER DEFINIZIONE

$$= \frac{P(X>t+s \cap X>t)}{P(X>t)} = \frac{P(X>t+s)}{P(X>t)}$$

$$= \frac{1-P(X<t+s)}{1-P(X<t)} = \frac{\cancel{1}-(\cancel{1}-e^{-\lambda(t+s)})}{\cancel{1}-(\cancel{1}-e^{-\lambda t})} = \frac{\cancel{e^{-\lambda t}} \cdot e^{-\lambda s}}{\cancel{e^{-\lambda t}}} = P(X>s)$$