DESCRIZIONE DISTRIBUZIONI

[DISCRETE]

- BERNOULLIANA = probabilità su due soli valori $(0,1)$ detti anche FALLIMENTO e SUCCESSO
- BINOMIALE = probabilità che descrive il numero di successi in un processo di Bernoulli, ovvero $S_n = X_1 + X_2 + X_3 + \dots + X_n$ che quindi somma n variabili INDIPENDENTI di uguale distribuzione Bernoulliana
  [ES: lanci moneta - estrazioni urna] → successo: p - fallimento: 1-p
- GEOMETRICA = probabilità su numeri naturali (compreso lo 0) che segue una progressione geometrica, ovvero la probabilità del primo successo richiede l'esecuzione di k prove indipendenti per essere verificato, quindi la probabilità di ottenere il primo successo è $(1-p)^{k-1}p$, mentre di ottenere un primo insuccesso è $(1-p)^k p$
- IPERGEOMETRICA = probabilità che descrive ad esempio estrazioni di palline senza reinserimento (perdenti/vincenti) da un'urna
  NB: la probabilità con reinserimento usa distribuzione binomiale
- POISSON = probabilità che descrive l'accadimento di eventi successivamente ed indipendentemente in un dato intervallo di tempo sapendo che mediamente se ne verifica un numero λ
  [ES. numero chiamate ricevute da un centralino in un determinato arco di tempo]

[CONTINUE]

- UNIFORME = è una distribuzione che è uniforme su un insieme, cioè che attribuisce la stessa probabilità a due intervalli della stessa lunghezza contenuti nell'insieme
- NORMALE = la somma di n variabili casuali con E(x) e VAR(x) finite tende ad una distribuzione normale. La normale dipende da due parametri, la media (μ) e la varianza (σ^2)
- ESPONENZIALE = è una distribuzione che descrive la "durata della vita" di un fenomeno che non "invecchia" (privo di memoria).
  [ES: durata della vita di una particella - durata richiesta di servizio]
- GAMMA = distribuzione che descrive la distribuzione esponenziale, ma viene anche utilizzata come modello per i tempi d'attesa nel teorema delle code

VAR DISCRETE

Parleremo di VAR casuali discrete quando l'insieme dei valori sarà finito, cioè avremo un numero finito di probabilità e quindi di eventi diversi

ES: LANCIO DI UNA MONETA : $T = 1/2$; $C = 1/2$ Numero discreto $= 2$

DISTRIBUZIONE	PARAMETRI	DENSITÀ/PROBABILITÀ	E(x)	VAR(x)	MOMENTI
• Bernoulli	$\sim Be(p)$ $p \in [0,1]$	$P(0) = 1-p$ $P(1) = p$	p	$p(1-p)$	$q + pe^t$
• Geometrica	$\sim Ge(p)$ $p \in]0,1[$	$p(1-p)^k$	$\frac{1}{p} - 2$	$(1-p)/p^2$	$pe^t/(1-qe^t)$
• Binomiale	$\sim Bi(n,p)$ $p \in [0,1]$	$\binom{n}{k}(p)^k(1-p)^{n-k}$	np	$np(1-p)$	$(q+pe^t)^n$
• ipergeometrica	$\sim Ip(m,N,k)$	$\frac{\binom{h}{k}\binom{m-h}{n-k}}{\binom{m}{n}}$	nh/m	$\frac{n(m-n)h(m-h)}{m^2(m-1)}$	/
• Poisson	$\sim P(\lambda)$	$\frac{\lambda^x}{x!} \cdot e^{-\lambda}$	λ	λ	$e^{\lambda(e^t-1)}$

VAR CONTINUE

Sono quelle variabili che possono assumere un insieme continuo di valori, per le quali l'insieme dei possibili valori ha cardinalità al più numerabile

ES: probabilità che la lunghezza di un foglio sia fra 3 e 4 centimetri (ci sono però ∞ valori fra 3 e 4), quindi si ricorre alla densità, dunque per trovare $P(3 <$ lunghezza foglio $< 4)$ si fa $\int_3^4 f(x)\,dx$

DISTRIBUZIONE	PARAMETRI	DENSITÀ	E(x)	VAR(x)	MOMENTI	RIPARTIZIONE
• UNIFORME	$\sim UNIF(a,b)$	$1/(b-a)$	$(b+a)/2$	$(b-a)^2/12$	$(e^{bt} - e^{at})/(bt-at)$	$(x-a)/(b-a)$
• GAMMA	$\sim Ga(\alpha,\lambda)$	$x^{k-1} \cdot e^{-x/\theta}/\theta^k\Gamma(k)$	$k\theta$	$k\theta^2$	$(1-\theta t)^{-k}$ $t < \theta^{-1}$	$\gamma(k, x\theta)/\Gamma(k)$
• ESPONENZIALE	$\sim Exp(\lambda)$	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$	$\left(1-\frac{t}{\lambda}\right)^{-1}$	$1 - e^{-\lambda x}$
• NORMALE	$\sim N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2	$e^{\left(\mu x + \frac{\sigma^2 x^2}{2}\right)}$	$[1 + ERF\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)]$

$$ERF = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}$$

DEFINIZIONI

- **FUNZIONE RIPARTIZIONE** [FUNZIONE DISTRIBUZIONE CUMULATIVA] = è una funzione di variabili reali che racchiude le informazioni su un fenomeno (insieme di dati, eventi casuali) riguardanti la sua distribuzione prima e dopo un certo punto, quindi la f.rip di una variabile casuale X a valori reali è la funzione che associa a ciascun valore x la probabilità dell'evento, assumendo valori minori/uguali x [$F(x) = P(X \le x)$]
  Si verifica con $\lim_{x \to \infty} F(x) = 1$; $\lim_{x \to -\infty} F(x) = 0$
  [DISCRETE] $F(x) = \sum_{x_i \le x} p(x_i)$ dove $p(x) = P(X = x)$ — es: lancio dadi $F(x) = \begin{cases} 0 & x<1 \\ x/6 & 1 \le x < 6 \\ 1 & x \ge 6 \end{cases}$
  [CONTINUE] $F(x) = \int_{-\infty}^{x} f(u)\,du$ dove $f(u)$ densità — es: è continua in $(0,1)$ $F(x) = \begin{cases} 0 & x<0 \\ x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$

- **INDIPENDENZA** [STOCASTICA] = L'indipendenza di due eventi A e B si ha quando il verificarsi di uno non modifica la probabilità di verificarsi dell'altro, più specificatamente quando $P(A|B)$ oppure $P(B|A)$ è pari rispettivamente a $P(A)$ e $P(B)$
  [FUNZIONE DI RIPARTIZIONE] Da $F_{XY}(x,y)$ trovo le due marginali $f_X(x) \to \int_a^b F_{XY}(x,y)\,dy$ e $f_Y(y) \to \int_a^b F_{XY}(x,y)\,dx$, se verificando $F_{XY}(x,y) = f_X(x) \cdot f_Y(y)$ allora sono indipendenti
  [FUNZIONE DENSITÀ] posso trovare le funzioni di probabilità/densità marginali facendo $f_X(x) \to \int_a^b f_{XY}(x,y)\,dy$ e $f_Y(y) \to \int_a^b f_{XY}(x,y)\,dx$ e verifico se $f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$

CONVERGENZA

- **TEOREMA LIMITE CENTRALE** [TLC] = la somma (normalizzata) di un gran numero di variabili casuali è distribuita approssimativamente con una variabile casuale normale standardizzata. Più tecnicamente il teorema afferma che se si ha una somma di variabili aleatorie X_i indipendenti e identicamente distribuite (iid) con media μ e varianza σ^2, allora indipendentemente dalla distribuzione di partenza, la somma tende a distribuirsi come una normale [TLC $\to P(X \le x) \to \frac{x-\mu}{\sigma} \cdot \sqrt{n}$ oppure $x-\mu/(\sigma/\sqrt{n})$]

- **LEGGE GRANDI NUMERI** [LGN] = descrive il comportamento della media di una sequenza di n variabili casuali indipendenti e con stessa distribuzione (n misure stessa grandezza; n lanci; ...) che si avvicina sufficientemente al reale valore di media
  [LEGGE DEBOLE] Data una successione di variabili casuali $X_1, X_2, X_3, \dots, X_n$ aventi stessa μ, σ e indipendenti si considera la media campionaria $\bar{X}_n = (X_1+X_2+X_3+\dots+X_n)/n$, la legge debole afferma quindi che per ogni $\varepsilon > 0$ $\lim_{n \to \infty} P(|\bar{X}_n - \mu| < \varepsilon) = 1$ ossia la $\bar{X}_n$ converge in probabilità alla media comune delle X_i

- **TEOREMA BAYES** = viene impiegato per calcolare la probabilità di una causa che ha provocato l'evento verificato. Ricordandoci il teorema moltiplicativo dove $P(E_2) \cdot P(E_1|E_2) = P(E_1) \cdot P(E_2|E_1)$ allora posso dire che $P(E_1|E_2)$ [probabilità che accada il primo visto che il secondo è già avvenuto] è uguale a $(P(E_2|E_1) \cdot P(E_1)) / P(E_2)$
  Ad esempio: probabilità che una persona soffra di una malattia (test positivo) o non soffra (test negativo) conoscendo la percentuale di veridicità del test e la frequenza con cui si presenta la malattia

- **TEOREMA MOLTIPLICATIVO** = verifica che il prodotto di un evento è uguale al prodotto fra la probabilità del primo evento e la probabilità del secondo condizionato dal primo già avvenuto
  $P(E_1 \cap E_2) \Rightarrow P(E_1)P(E_2|E_1) = P(E_2)P(E_1|E_2)$

- **PROBABILITÀ TOTALE** = consente di calcolare la probabilità che si verifichi almeno uno di due o più eventi cioè la probabilità dell'unione di essi. Ha due formulazioni: per eventi incompatibili o eventi qualsiasi
  [INCOMPATIBILI] Sono incompatibili quando solo uno dei due può avvenire, quindi $P(E) = P(E_1) \cdot P(E|E_1) + P(E_2)P(E|E_2)$ si nota che l'evento E può occorrere solo se associato ai due precedenti. Dimostriamo $\to E = (E \cap E_1) \cup (E \cap E_2)$ che per la proprietà additiva fra eventi incompatibili $P(E) = P(E \cap E_1) + P(E \cap E_2)$, quindi per il teorema moltiplicativo $P(E) = P(E_1)P(E|E_1) + P(E_2)P(E|E_2)$ cvd (es: lancio ~~dado ottengo~~ '1' e '6' $P(X>4) \to \frac{1}{6} + \frac{1}{6} = 1/3$)
  [EVENTI QUALSIASI] Se non sono incompatibili, bisogna tenere conto delle loro intersezioni. In generale la probabilità di due eventi A e B è uguale alla somma delle singole probabilità $(P(A)+P(B))$ − la probabilità delle loro intersezioni $(P(A \cap B))$
  (es: lancio due dadi $A_1 = '6'$, $A_2 = '6'$; l'evento "almeno un dado dia 6" è l'unione dei due eventi non incompatibili $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \to 1/6 + 1/6 - 1/36 = 11/36$)

- $\boxed{\text{FUNZIONE (DISTRIBUZIONE) MARGINALE}}$ = Date due variabili casuali X e Y la cui distribuzione congiunta sia nota, la distribuzione marginale di X è semplicemente la distribuzione di probabilità di X mediante sopra l'informazione di Y

[DISCRETE] Può essere scritta come $Pr(X=x, Y=y)$ se è una distribuzione congiunta ; $Pr(X=x \mid Y=y)$ se è una distribuzione con X condizionato Y $\left[Pr(X=x) = \sum_y Pr(X=x, Y=y) = \sum_y Pr(X=x \mid Y=y)\, Pr(Y=y) \right]$

[CONTINUE] Può essere scritta come $P_X(x) = \int_y P_{X,Y}(x,y)\, dy = \int_y P_{X|Y}(x|y)\, P_Y(y)\, dy$, dove $P_{X,Y}(x,y)$ dà la distribuzione congiunta di X e Y , mentre $P_{X|Y}(x|y)$ fornisce la distribuzione condizionata per X dato Y

- $\boxed{\text{FUNZIONE GENERATRICE MOMENTI}}$ (FGM) = Permette di estrarre agevolmente alcuni parametri come valore atteso [E(x)] e varianza [VAR(x)], o anche confrontare due variabili casuali diverse vedendo il loro comportamento. La FGM [$g(t)$] di una variabile casuale X è definita come il valore atteso di e^{tX} dove esso è finito (solo in un intorno di 0 in cui vale 1 indipendentemente da x). Se e^{tx} fosse infinito si dice semplicemente che X non ha FGM

[DISCRETE] $g(t) = E(e^{tX}) = \sum_{i=1}^{n} p_i e^{tx_i}$

[CONTINUE] $g(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \cdot f_X(x)\, dx$ (densità)

[TEOREMI] Se $X_1, X_2, \dots, X_n$ sono variabili casuali indipendenti e X la loro somma, allora la FGM di X è il prodotto delle funzioni generatrici dei momenti delle singole X_i $\quad g(t;X) = \prod_{i=1}^{n} g(t; X_i)$

- $\boxed{\text{VALORE ATTESO}}$ [E(x)] = In una VARIABILE DISCRETA (cioè che assume solo un numero finito o una infinità numerabile di valori) è dato dalla somma dei possibili valori di tale variabile, ciascuno moltiplicato per la probabilità di essere assunto (= verificarsi) $\left[E(x) = \sum_{i=1}^{\infty} x_i p_i \right]$ (es: Testa o croce, ipotizziamo 100 T $\to E(x) = \frac{1}{2} \cdot 100 + \frac{1}{2} \cdot 0$ / C 100)

Per una VARIABILE CONTINUA che ammette una funzione di densità [$f(x)$] basterà integrare quest'ultima, nell'intervallo in cui è definita, moltiplicandola per x $\to \int_{-\infty}^{+\infty} x\, f(x)\, dx$

- $\boxed{\text{VARIANZA}}$ [VAR(x)] = è una funzione indicata con $\sigma^2(x)$ che fornisce una misura di quanto siano vari i valori assunti dalla variabile, cioè quanto si discostino dal valore atteso. Si calcola in questo modo $\to VAR(x) = E(x^2) - E(x)^2$

[PROPRIETÀ] Non è mai negativa ed è 0 solo quando $P(X=x)=1$; la varianza di due variabili indipendenti è pari alla somma delle loro varianze $[\sigma^2(x+y) = \sigma^2(x) + \sigma^2(y)]$; nel caso in cui avremmo $\sigma^2(x-y)$ $\sigma^2(x+(-y)) \to$ $\to \sigma^2 + \sigma^2(-y) \to \sigma^2(x) + \sigma^2(y)$; se X e Y non fossero indipendenti $\sigma^2(x+y) = \sigma^2(x) + \sigma^2(y) + 2\sigma^2(x,y)^{\leftarrow COV}$; $\sigma^2(2x+3y) = 4\sigma^2(x) + 9\sigma^2(y)$

[DISCRETE] Se $E(x) = \sum_{y \in X} y P(X=y) \to \sigma^2(x) = \sum_{y \in X} (y - E(x))^2 P(X=y)$ $\quad$ NB: $VAR(x) = COV(X,X)$

[CONTINUE] Se $E(x) = \int_x y f(x)\, dy \to \sigma^2(x) = \int_x (y - E(x))^2 f(y)\, dy$

- $\boxed{\text{COVARIANZA}}$ [$COV(X,Y) \Leftrightarrow \overset{\sigma_{X,Y}}{\sigma^2(X,Y)}$] = è una misurazione di quanto X e Y variano insieme, ovvero della loro DIPENDENZA. Può essere definita come la differenza tra il valore atteso del loro prodotto e il prodotto dei loro valori attesi $COV(X,Y) = E(X,Y) - E(X) \cdot E(Y)$

[PROPRIETÀ] • $COV(X,Y) = COV(Y,X)$; • $COV(aX+b, Y) = a\,COV(X,Y)$; • $COV(X+Y, Z) = COV(X,Z) + COV(Y,Z)$;
• se $A \perp\!\!\!\perp B \to COV(X,Y) = 0$; • Dipendenti ma non correlate $(X \sim UNIF(-1,1)$ e $Y = X^2)$ allora $\longrightarrow$
$\to COV(X,Y) = COV(X,X^2) = E(X^3) - E(X)E(X^2) = 0 - 0E(X^2) = 0$

- $\boxed{\text{CORRELAZIONE}}$ [CORR(X,Y)] = la correlazione fra due variabili X e Y è il rapporto tra la covarianza e il prodotto delle loro deviazioni standard (VARIANZA) $\to CORR(X,Y) = COV(X,Y) / (\sqrt{VAR(X)} + \sqrt{VAR(Y)})$

[PROPRIETÀ] • $CORR(X,X) = COV(X,X) / (\sigma(x) \cdot \sigma(x)) \to E((x-\bar{x})(x-\bar{x})) / \sigma^2(x) \to VAR(x)/VAR(x) \to 1$
• $CORR(X,-X) = COV(X,-X) / (\sigma(x) \cdot \sigma(-x)) \to E((x-\bar{x})(-x-(-\bar{x}))) / \sigma^2(x) \to -VAR(x)/VAR(x) \to -1$
• Se $X \perp\!\!\!\perp Y \to CORR(X,Y) = E(XY) - \bar{X}\bar{Y} / \sigma(x)\sigma(y) \to \bar{X}\bar{Y} - \bar{X}\bar{Y} / \sigma^2(x) \to 0$

RIEPILOGO FORMULE : VAR ALEATORIE CONTINUE

- $f_X(x) = k e^{-|x|}$
  - $\int_{-\infty}^{+\infty} k e^{-|x|} = 1 \to k \int_{-\infty}^{+\infty} e^{-|x|} = 1 \to k\,[e^{x}]_{-\infty}^{0} + [-e^{x}]_{0}^{\infty} = 1 \to k\,[1-0] + [0+1] \to 2k = 1 \to k = 1/2$
  - Densità $\to$ funzione Ripartizione $[\int_a^b]$
  - funzione Ripartizione $\to$ Densità [Derivata]

- $F_Y(y) = 1 - 1/y^2 \qquad 1 < y < \infty$
  - verificare se densità : $\lim_{x \to 1} F_Y(y) = \ell \quad ; \quad \lim_{x \to \infty} F_Y(y) = \ell$ verificata se $\ell \neq \infty$
  - $f_Y(y) = F_Y(y)' \to 2/y^3$
  - MODA = $F_Y(y) = 1 \to 1 - \frac{1}{y^2} = 1 \to y = 1$
  - MEDIANA: $F_Y(y) = 0.5 \to 1 - 1/y^2 = 0.5 \to x = \sqrt{2}$
  - Fiume supera la soglia di guardia di $k > 1 \to P(Y > k) \to 1 - P(Y < k) \to 1 - F_Y(k) \to 1 - (1 - 1/k^2) \to 1/k^2$

- $f_X(x) = c/x^2 \qquad \mathbb{1}_{(100,\infty)}(x)$
  - valori di $c \to \int_{100}^{\infty} \frac{c}{x^2} = 1 \to c \int_{100}^{\infty} \frac{1}{x^2} \to c\,[-\frac{1}{x}]_{100}^{\infty} = 1 \to c/100 = 1 \to c = 100$
  - $F_X(x) = \int_{100}^{x} \frac{100}{x^2} \to 100 \int_{100}^{x} \frac{1}{x^2} \to 100\,[-\frac{1}{x}]_{100}^{x} \to 100\,(-\frac{1}{x} + \frac{1}{100}) \to 1 - \frac{100}{x}$
  - QUARTILI $\to q_\alpha = (F_X(x) = \alpha) \to q_{0.25}: 1 - 100/x = 0.25 = \frac{400}{3}$; $q_{0.5}$ (mediana) : $1 - \frac{100}{x} = 0.5 = 200$; $q_{0.75}: 1 + \frac{100}{x} = 0.75 = 400$
  - $P(X > 500) \to 1 - P(X < 500) \to 1 - (F_X(x) \to 1 - (1 - \frac{100}{500}) \to 1/5$

- $Z \sim N(0,1)$
  - $P(Z < a) = 0.9686 \to \Phi(a) = 0.9686 \to a = 1.86$
  - $P(Z > b) = 0.1788 \to 1 - P(Z < b) = 0.1788 \to P(Z < b) = 0.8212 \to \Phi(b) = 0.8212 \to b = 0.92$

- $X \sim N(3, 5^2) \equiv X \sim N(\mu, \sigma^2) \qquad P(a < x < b) \to P(\frac{a-\mu}{\sigma} < x < \frac{b-\mu}{\sigma})$
  - $P(4 < x < 6) \to P(\frac{4-3}{5} < x < \frac{6-3}{5}) \to \Phi(3/5) - \Phi(1/5) \to 0.72575 - 0.579 = 0.14649$
  - $P(1 < x < 5) \to P(\frac{1-3}{5} < x < \frac{5-3}{5}) \to \Phi(2/5) - (1 - \Phi(2/5)) \to 0.655 - (1 - 0.655) = 0.3108$
  - $P(-1 < x < 2) \to P(\frac{-1-3}{5} < x < \frac{2-3}{5}) \to (1 - \Phi(4/5)) - (1 - \Phi(1/5)) \to 0.78814 - 0.5792 = 0.2088$

- $X \sim N(4, 4^2) \qquad c = ?$
  - $P(|x-4| \leq c) = 0.95 \to P(-c < x-4 < c) = 0.95 \to P(4-c < x < 4+c) = 0.95 \to P(\frac{4-c-4}{4} < x < \frac{4+c-4}{4}) = 0.95$
    $\to P(-c/4 < x < c/4) = 0.95 \to 2P(x < c/4) - 1 = 0.95 \to P(x < c/4) = \frac{1+0.95}{2} \to \Phi(\frac{c}{4}) = 0.975 \to \frac{c}{4} = 1.96 \to c = 7.84$

- $X \sim N(\mu, \sigma^2)$ $\mu = ?$ $\sigma^2 = ?$
  - $P(X \le 2.45) = 0.15 \to P\left(X \le \frac{2.45-\mu}{\sigma}\right) = 0.15 \to q_{0.15} = -q_{0.85} \to \Phi\left(\frac{2.45-\mu}{\sigma}\right) = -0.85 \to \frac{2.45-\mu}{\sigma} = -1.04$
  - $P(X \ge 2.6) = 0.06 \to 1-P\left(X < \frac{2.6-\mu}{\sigma}\right) = 0.06 \to P\left(X < \frac{2.6-\mu}{\sigma}\right) = 0.94 \to \Phi\left(\frac{2.6-\mu}{\sigma}\right) = 0.94 \to \frac{2.6-\mu}{\sigma} = 1.56$
  - $\begin{cases} \frac{2.45-\mu}{\sigma} = -1.04 \\ \frac{2.6-\mu}{\sigma} = 1.56 \end{cases} \to \begin{cases} \mu = 2.51 \\ \sigma = 0.057 \end{cases} \to \begin{cases} \mu = 2.51 \\ \sigma^2 = 0.0033 \end{cases}$

- $X \sim \text{UNIF}(a,b)$ $E(X) = 6 = (a+b)/2$ $VAR(X) = 2 = (b-a)^2/12$
  - $\begin{cases} a+b/2 = 6 \\ (b-a)^2/12 = 2 \end{cases} \Rightarrow \begin{cases} a = 3.55 \\ b = 8.45 \end{cases}$
  - $f_X(x) = 1/(b-a) \to 1/4.9$
  - $F_X(x) = (x-a)/(b-a) \to (x-3.55)/(4.9)$
  - MEDIANA: $F_X(x) = 0.5 \to \frac{x-3.55}{4.9} = 0.5 \to x = 7.225$
  - $P(X < 5.5) \to \int_a^{5.5} f_X(x) \to \frac{1}{4.9}\int_a^{5.5} 1 \to \frac{1}{4.9}\cdot[x]_{3.55}^{5.5} \to \frac{1}{4.9}\cdot(5.5-3.55) \to 1.95/4.9$

- $X \sim N(20, \sigma^2)$ $q_{0.3}$ = supero esame @[VOTO] superato da $q_{0.10} = ?$
  - $P(X \le 18) = 0.3 \to q_{0.3} = -q_{0.7} \to P(X = 18-20/\sigma) = -0.7 \to \Phi\left(\frac{-2}{\sigma}\right) = -0.7 \to -2/\sigma = -0.52 \to \sigma = 3.846$
  - $P(X>a) = 0.1 \to 1-P(X<a) = 0.1 \to 1-P\left(X < \frac{a-20}{3.846}\right) = 0.1 \to P(X < a-20/3.846) = 0.9 \to \Phi\left(\frac{a-20}{3.846}\right) = 0.9 \to \frac{a-20}{3.846} = 1.28 \to a = 24.92$

- $X \sim \text{UNIF}(2,6)$ $Y \sim \text{Exp}(1/h)$
  $a^x = b \iff x = \log_a b$
  - $P(X < 4) \to \frac{x-a}{b-a} \to \frac{4-2}{6-2} \to 1/2$
  - $P(Y < 4) \to 1-e^{-\lambda x} \to 1-e^{-1/h \cdot x} \to 1-e^{-4/h}$
  - $e^{-4h} = 1/2 \to -4h = \ln_e 1/2 \to h = \ln 1/2 / -4 \to 0.1732$ $h = 0.1732$

- $X \sim \text{Exp}(\lambda)$ $F_X(x) = 1-e^{-7x}$
  - Durata di 3 mesi dato che è già durata 1 mese: $P(X>3 \mid X>1) \to P(X>2) \to 1-P(X<2) \to 1-(1-e^{-7(2)}) \to e^{-14}$
  - Su 30 batterie 5 durino più di 1/3 di mese ; $Y \sim Bi(30, e^{-7(1/3)})$ $P(X>1/3) \to 1-P(X<1/3) \to e^{-7/3}$
  $\to P(X<5) \to \sum_{i=0}^{5} \binom{30}{i} (e^{-7/3})^i (1-e^{-7/3})^{30-i} \to 0.8398$

- $X \sim N(\mu, \sigma^2)$ $\mu = 300$ $\sigma^2 = ?$
  - P che un barattolo contenga 304 è pari a 0,1: $P(X>304) = 0.1 \to 1-P(X<304) = 0.1 \to P(X<304) = 0.9 \to P\left(X < \frac{304-300}{\sigma}\right) = 0.9$
  $\to \Phi\left(\frac{4}{\sigma}\right) = 0.9 \to 4/\sigma = 1.29 \to \sigma = 3.1007$ $\sigma^2 = 9.61$
  - Barattoli conformi se contengono 297: $P(X>297) \to 1-P(X<297)$ NON CONFORMI $\to 1-P\left(X < \frac{297-300}{3.1007}\right) \to 1-(1-\Phi(0.96)) \to 0.8315$
  - P che su 25 barattoli, più di 2 non conformi: $X \sim Bi(25, (1-0.8315))$, $P(X>2) \to 1-P(X<2) \to 1-\left(\sum_{i=0}^{2} \binom{25}{i} (0.8315)^i (1-0.8315)^{25-i}\right) \to 0.877$

- $X \sim U(a,b) \to (0,1)$

  - $f_X(x) = 1/(b-a) \to 1 \quad \begin{cases} 1 & 0<x<1 \\ 0 & \text{else} \end{cases}$
  - $F_X(x) = (x-a)/(b-a) \to x \quad \begin{cases} x & 0<x<1 \\ 0 & \text{else} \end{cases}$

  - $E(x) = \int_a^b x/(b-a) \to (a+b)/2 \to 1/2$
  - $E(x^2) = \int_a^b x^2/(b-a) \to (b+a)^2/3 \to 1/3$
  - $VAR(x) \to (b-a)/12 \to 1/12$

  - $P(X > z + E(x)) = 1/4 \to 1 - P(x < z + E(x)) \to P(x < z + E(x)) = 3/4 \to \frac{1}{a-b}\int_0^{z+1/2} 1 \to 1[x]_0^{z+1/2} \to z+1/2-0 = 3/4 \to z = 1/4$

- $X \sim Exp(h) \to (0.005)$

  - lampadina dura almeno 300 giorni : $P(x \geq 300) \to e^{-hx} \to e^{-0.005 \cdot 300} \to 0.223$
  - Dopo 150 giorni duri ancora 120 giorni : $P(x > 150+120 \mid x > 150) \to P(x > 120) \to e^{-0.005 \cdot 120} \to 0.5488$
  - Prob che su 50 lampadine solo 3 durino più 300 giorni : $X \sim Bi(50, 0.5488) \to P(x=3) \to \binom{50}{3}(0.223)^3(1-0.223)^{50-3} \to 0.00153$

- $X \sim N(58.5, 0.3)$

  - $P(58.1 < x < 58.8) \to P\left(\frac{58.1-58.5}{\sqrt{0.3}} < x < \frac{58.8-58.5}{\sqrt{0.3}}\right) \to \Phi(0.3162) - (1-\Phi(0.4216)) \to 0.2874$
  - Diametro 58.7 visto che supera la media : $P(x > 58.7 \mid x > 58.5) \to [P(X > 58 \mid x > 58.5) P(x > 58.7)] / P(x > 58.5)$

    $\to P(x > 58.7)/P(x > 58.5) \to 1 - P\left(x < \frac{58.7-58.5}{\sqrt{0.3}}\right) / 1 - P\left(\frac{58.5-58.5}{\sqrt{0.3}}\right) \to 1 - \Phi(0.2168)/1-\Phi(0) \to 0.8330$
  - Su 10 ruote solo 1 con diametro < 58.3 : $P(x < 58.3) \to P\left(X < \frac{58.3-58.5}{\sqrt{0.3}}\right) \to 1 - \Phi(0.2288) \to 0.41651$

    $\to X \sim B(10, 0.41651) \to y = 1$ ruote $\to \binom{10}{1}(0.41651)^1(1-0.41651)^{10-1} \to 0.3265$
  - Su 20 ruote 2 hanno $d < 58.3 \to 1/\binom{20}{2} \to 0.0053$

- lunghezza$(A) \sim UNIF(10-a, 10+a)$ lunghezza$(b) \sim N(10, 4)$ $P(A) = 2/3 \iff P(B) = 1/3$

  - Pezzo fuori dalle specifiche $(9,11)$ (U) $P(X \neq (9,11) \mid A) \to 1 - P(X = (9,11) \mid A) \to 1 - P(9 \mid A < x < 11 \mid A)$

    $\to 1 - P\left(\frac{9-(10-a)}{2a} < x < \frac{11-(10-a)}{2a}\right) \to 1 - P\left(\frac{x-a}{b-a} < x < \frac{x-a}{b-a}\right) \to 1 - 1/a$

    $\to (N) \to P(x \neq (9,11) \mid B) \to 1 - P(x = (9,11) \mid B) \to 1 - P\left(\frac{a-\mu}{\sigma} < x < \frac{b+\mu}{\sigma}\right) \to 1 - P\left(\frac{9-10}{2} < x < \frac{11-10}{2}\right)$

    $\to 1 - P(-1/2 < x < 1/2) \to 1 - (\Phi(1/2) - (1-\Phi(1/2))) \to 0.6172$

    $\to (TOT) \to P(x \neq (9,11) \mid A) \cdot P(A) + P(x \neq (9,11) \mid B) P(B) \to \frac{2}{3}(1 - \frac{1}{a}) + \frac{1}{3}(0.6172) \to 0.872 - 2/3a$

  - $E(X \mid A) \to (a+b)/2 \to (10-a+10+a)/2 \to 10$
  - $E(X \mid B) = 10$
  - $VAR(X \mid A) \to (b-a)^2/12 \to (10+a-10+a)^2/12 \to a^2/3$
  - $VAR(x \mid B) = 4$
  - $\frac{a^2}{3} = 4 \to a = 2\sqrt{3}$

- $X \sim Exp(\frac{1}{2}) \quad f_X(x) = \frac{1}{2} \cdot e^{-x/2}$

  - Su 100 componenti, k durano più di 2 ore : $Y \sim Bi(100, e^{-(2)/2}) \to P(x > 2) \to \int_2^{\infty} 1/2 \cdot e^{-x/2}$

    $\to [-e^{-x/2}]_2^{\infty} \to e^{-1}$

    $\to P(Y = k) \to \binom{100}{k}(e^{-1})^k(1-e^{-1})^{100-k}$

RIEPILOGO FORMULE : TRASFORMAZIONE VAR ALEATORIE

- Almeno 50 persone favorevoli: $P(X > 49.5) \to 1 - P(X \leq 49.5) \to 1 - P(Z \leq 49.5 - 65 / 22.75)$
- tra 60 e 70 persone favorevoli: $P(59.5 < X < 70.5) \to P(59.5 - 65/22.75 < x < 70.5 - 65/22.75)$
- meno di 75 persone favorevoli: $P(X < 74.5) \to P(Z \leq 74.5 - 65/22.75)$
- $X \sim Bi(m, p) \equiv X \sim N(mp, mp(1-p))$

- $M_X(t) = e^{2t + 2t^2}$
  - $M_X'(t) = (2 + 4t) e^{(2t+2t^2)}$ $\qquad M'(0) = E(X) = 2$
  - $M_X''(t) = (2+4t)^2 \cdot e^{(2t+2t^2)} \cdot 4 e^{(2t+2t^2)}$ $\qquad M''(0) = E(X^2) = 8$
  - $2t + 2t^2 = \mu t + \frac{t^2 \sigma^2}{2}$ $\qquad \mu = 2 \quad \sigma^2 = 4$
  - $P(0 < X < 4) \Rightarrow P\left(\frac{0-2}{\sqrt{4}} < x < \frac{4-2}{\sqrt{4}}\right)$

- $X \sim UNIF(-1, 1)$ $\qquad f_n = (x - \alpha)/(\beta - \alpha)$
  - $P(|X| > 1/2) \to P(-1/2 < x < 1/2) \to P\left(-\frac{1}{2}+1/_{1+1} < x < \frac{1}{2}+1/_{1+1}\right) \to 0.75 - 0.25 \to 0.5$
  - Densità $|X| \to f_Y(y) = P(y < Y) \to P(|X| < Y) \to P(-Y < X < Y) \to P(-Y+1/_{1+1} < x < Y+1/_{1+1}) \to Y$
    $F_Y(y) \begin{cases} y & 0<y<1 \\ 0 & \text{else} \end{cases}$ $\qquad f_Y(y) \begin{cases} 1 & 0<y<1 \\ 0 & \text{else} \end{cases}$

- $X \sim Exp(\lambda)$ $\qquad$ distrib? $Y = F_X(x)$ $\qquad$ rip $Y \equiv$ densità X
  - densità $X \to f_X(x) = \lambda e^{-\lambda x}$ $\qquad$ • ripartiz. y $[F_Y(y)] = \lambda e^{-\lambda x} \equiv F_X(x)$
  - $g^{-1}(y) = -\ln(1-y)/\lambda$ $\qquad$ • $Dg^{-1}(y)/Dy = 1/\lambda(1-y)$
  - Densità y $[f_Y(y)] = (f_X(x) \circ g^{-1}) \cdot \frac{Dg^{-1}(y)}{Dy} \to \lambda e^{-\lambda(-\ln(1-y)/\lambda)} \cdot \frac{1}{\lambda(1-y)} \to f_Y(y) = 1$
  - $f_Y(y) = 1$ $\qquad \mathbb{1}_{(0,1)}(y)$ $\qquad Y \sim UNIF(0,1)$

- $f_X(x) = X^2/9$ $\qquad 0 < x < 3$ $\qquad f_Y(y)?$ $Y = X^3$
  - $F_Y(y) \equiv f_X(x) \to \frac{x^2}{9}$ $\qquad$ • $g^{-1}(y) = y^{1/3}$ $\qquad$ • $\frac{Dg^{-1}(y)}{Dy} = \frac{1}{3} y^{-2/3}$
  - $f_Y(y) = (f_X(x) \cdot g^{-1}(y)) \cdot \frac{Dg^{-1}(y)}{Dy} \to \frac{(y^{1/3})^2}{9} \cdot \frac{1}{3} y^{-2/3}$

- $f_X(x)(x, \lambda) = \lambda e^{-\lambda x}$ $\qquad$ densità? $Y = \ln X$ $\qquad$ densità? $Z = X/(X+1)$
  - $F_Y(y) = \ln x$ $\qquad$ • $g^{-1}(y) = e^y$ $\qquad$ • $\frac{Dg^{-1}(y)}{Dy} = e^y$ $\qquad$ • $f_Y(y) = \lambda e^{-\lambda e^y} \cdot e^y$ $\qquad (-\infty, \infty)$
  - $F_Z(z) = X/(x+1)$ $\qquad$ • $g^{-1}(z) = z/1-z$ $\qquad$ • $\frac{Dg^{-1}(z)}{Dz} = 1/(1-z)^2$ $\qquad$ • $f_Z(z) = \lambda e^{-\lambda(z/1-z)} \cdot 1/(1-z)^2$ $\qquad (0, 1)$

- $f_X(x) = 1/20 \sim U(0,20)$ μ_{ed}? $Y = (x+3)/2$
  - $F_Y(y) = (x+3)/2$
  - $g^{-1}(y) = 2y-3$
  - $\frac{Dg^{-1}(y)}{Dy} = 2$
  - $\frac{x+3}{2} \to \alpha = 0+3/2$, $\beta = 20+3/2$
  - $f_Y(y) = \begin{cases} 1/\beta-\alpha & 3/2 < y < 23/2 \\ 0 & \text{else} \end{cases}$
  - $E(Y) = \beta+\alpha/2 = 13/2$
  - $E((x+3)/2) = \frac{E(x)}{2} + \frac{3}{2} = 13/2$
  - $VAR(y) = (\beta-\alpha)^2/12 = 8.33$

- $X \sim Bi(200, 0.12) \equiv X \sim N(24, 88, 12)$
  - Almeno 20 moncini $= P(X \geq 19.5) \to 1-P(X < 19.5) \to 1-P(x < (m-np)/\sqrt{np(1-p)}) \to 1-P(x < (19.5-24)/\sqrt{21.12})$
    $\to 1-(1-I(0.88))$

- $M_X(t) = 9/(3-t)^2$ $\sigma^2, \mu = ?$
  - $M'_X(t) = -18/(t^3-9t^2+27t-27)$
  - $M'_X(0) = E(x) = 2/3$
  - $M''_X(t) = 54/(t^4-12t^3+54t^2-108t+81)$
  - $M''_X(0) = E(x^2) = 2/3$
  - $VAR(x) = E(x^2)-E(x)^2 = 2/9$

- $f_X(x) = \lambda e^{-\lambda x} \sim Exp(\lambda)$ distrib? $Y = X^2$
  - $F_Y(y) = x^2$
  - $g^{-1}(y) = y^{1/2}$
  - $\frac{Dg^{-1}(y)}{Dy} = \frac{1}{2} y^{-1/2}$
  - $f_Y(y) = \lambda e^{-\lambda y^{1/2}} \cdot \frac{1}{2} y^{-1/2}$ $(0, +\infty)$

- $f_X(x) = \lambda e^{-\lambda x} \sim Exp(\lambda)$ ε exp? $Y = cX$ $c > 0$
  - $F_Y(y) = P(Y \leq y) - P(cX < y) \to P(X < y/c) \to F_X(y/c)$
  - $1 * e^{-\lambda(y/c)} \to 1 * e^{-y(\lambda/c)} \to Y \sim Exp(\lambda/c)$

- $X \sim UNIF(0,1) \to f_X(x) = 1/(b-a) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{else} \end{cases}$ densità? $Y = e^X$ $\alpha = e^0$ $\beta = e^1$
  - $F_Y(y) = e^x$
  - $g^{-1}(y) = \ln y$
  - $\frac{Dg^{-1}(y)}{Dy} = 1/y$
  - $f_Y(y) = \begin{cases} 1/y & 1 < y < e \\ 0 & \text{else} \end{cases}$

- $X \sim N(5,4)$ $E(Y)$? $Y = (X-5)^2$
  - $E(Y) = E((y-5)^2) \to E((x-E(x))^2) \to VAR(x) = 4$ $VAR(Y) = 4$

- $X \sim GE(p) \to f_X(x) = p(1-p)^k$ $E(Y)$? prob(densità)? $Y = 2X-1$
  - $E(x) = 1/p$
  - $E(Y) \to E(2X-1) \to 2\frac{1}{p} - 1 = (2-p)/p$
  - $g^{-1}(y) = \frac{x}{2} + \frac{1}{2}$
  - $\frac{Dg^{-1}(y)}{Dy} = 1/2$
  - $P(Y=k) \to P(2X-1=k) \to P(X=(k+1)/2) \to (1-p)^{\frac{k+1}{2}} p \cdot 1/2$

RIEPILOGO FORMULE: DISTRIBUZIONI CONGIUNTE

- 3 palline numerate 1:3   2 SR   X = più grande numero estratto   Y = somma dei numeri estratti

  - 12, 13, 21, 23, 31, 32 → $X : \begin{cases} 2 & 3 \\ 1/3 & 2/3 \end{cases}$ $Y = \begin{cases} 3 & 4 & 5 \\ 2/6 & 2/6 & 2/6 \end{cases}$ <=> X {MAX NUM 2, MAX NUM 3}   Y {SOMMA 3, SOMMA 4, SOMMA 5}

  - funzione di probabilità $P(x,y) = P(X=x, Y=x)$

X \ Y	3	4	5	
2	1/3	0	0	1/3
3	0	1/3	1/3	2/3
	1/3	1/3	1/3	

MAX 2 + SOMMA 3 => 2/6   MAX 3 + SOMMA 3 = 0
MAX 2 + SOMMA 4 => 0   MAX 3 + SOMMA 4 = 2/6
MAX 2 + SOMMA 5 => 0   MAX 3 + SOMMA 5 = 2/6

  - Y condizionato X=3 → $P(Y|X=3) = P(Y=y|X=3)/P(X=3)$ → $\begin{cases} 3 & 4 & 5 \\ 0 & 1/2 & 1/2 \end{cases}$ → $\begin{cases} 0 & y<3 \\ 1/2 & 3 \le y < 4 \\ 1 & y \ge 4 \end{cases}$

  - $COV(X,Y) = E(XY) - E(X) \cdot E(Y) \to 11 - (\frac{8}{3})(4) = 1/3$ - $E(XY) = 2 \cdot 3 \cdot \frac{1}{3} + 2 \cdot 4 \cdot 0 + 2 \cdot 5 \cdot 0 + 3 \cdot 3 \cdot 0 + 3 \cdot 4 \cdot 1/3 + 3 \cdot 5 \cdot 1/3 = 11$

  - $E(X) = 2 \cdot \frac{1}{3} + 3 \cdot \frac{2}{3} = 8/3$ - $E(Y) = 3 \cdot \frac{1}{3} + 4 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} = 4$

  - $COR(X,Y) = COV(X,Y) / (\sqrt{VAR X} \cdot \sqrt{VAR Y}) = 0.3$ - $VAR(X) = E(X^2) = 4 \cdot \frac{1}{3} + 9 \cdot \frac{2}{3} = \frac{22}{3}$

  - $VAR(Y) = E(Y^2) = 9 \cdot \frac{1}{3} + 16 \cdot \frac{1}{3} + 25 \cdot \frac{1}{3} = 50/3$ - $E(XY) \neq E(X) \cdot E(Y)$

  - $E(2X+3Y) \to 2E(X) + 3E(Y) \to 52/3$ - $VAR(2X+3Y) + 2Cov(2X,3Y) \to 4VAR(X) + 9VAR(Y) + 2 \cdot 2 \cdot 3\, Cov(X,Y) = \frac{550}{3}$

- $P_{X,Y}(x,y) = k(2y+x)$ X = 2, 4   Y = 0, 1, 2

X \ Y	0	1	2	
2	2k	4k	6k	12k
4	4k	6k	8k	18k
	6k	10k	14k	30k

(2,0) → k(0+2) = 2k   (4,0) → k(0+4) = 4k
(2,1) → k(2+2) = 4k   (4,1) → k(2+4) = 6k
(2,2) → k(4+2) = 6k   (4,2) → k(4+4) = 8k

  - $30k = 1 \to k = 1/30$
  - $P(Y \ge X) \to P(X = \overset{MAX\,X}{\le Y} = 2, Y = \overset{MAX}{Y} = 2) \to 1/30\,(2 \cdot 2 + 2) \to 1/15$
  - $F_{X,Y}(2,1) \to P(X \le 2, Y \le 1) \to P(X=2, Y=0) + P(X=2, Y=1) \to \frac{1}{30}(0+2) + \frac{1}{30}(2+2) \to \frac{6}{30} \to 1/5$
  - $F_{XY}(4,1) \to \ldots \to P(X=4, Y=0) + P(X=4, Y=1) + P(X=2, Y=0) + P(X=2, Y=1) \to 8/15$
  - $P_{X|Y}(x|1) \to P(X=x, Y=1) \to P(x,1)/P_Y(1) \to 1/30\,(2+x) / P(X=2,Y=1) + P(X=4,Y=1) \to \frac{2+x/30}{4/30 + 6/30} \to 2+x/10$
  - $\perp \to P_X(x) \cdot P_Y(y) = P(x,y)$ - $P_X(x) \Rightarrow P(X=x, Y=0) + P(X=x, Y=1) + P(X=x, Y=2) \to \frac{2+x}{10}$
  - $\frac{2+x}{10} \cdot \frac{2y+3}{15} \neq \frac{2y+x}{30}$ NON DIPENDENTI   - $P_Y(y) \to P(X=2, Y=y) + P(X=4, Y=y) \to \frac{2y+3}{10}$

- [continui] $f_{XY}(x,y) = 12xy(1-y)$ $0<x<1$; $0<y<1$
  - [MARGINALE] $f_X(x) = \int_0^1 f_{XY}(x,y)\,dy \to \int_0^1 12xy(1-y)\,dy \to \int_0^1 12xy - 12xy^2 \to \int_0^1 6xy^2 - 4xy^3 \to 2x$
  - [MARGINALE] $f_Y(y) = \int_0^1 f_{XY}(x,y)\,dx \to \int_0^1 12xy(1-y)\,dy \to \int_0^1 12xy - 12xy^2 \to \int_0^1 6x^2y - 6x^2y^2 \to 6(y-y^2)$
  - $f_{XY}(x,y) = f_X(x) \cdot f_Y(y) \to 12xy(1-y) \equiv 2x \cdot (6y - 6y^2) \to$ sono INDIPENDENTI

- $f_{xy}(x,y) = k \qquad 0<x<1 \; ; \; 0<y<x$
  - [TROVARE K] $\int_0^1 \int_0^x f_{xy}(x,y)\,dx\,dy = 1 \to k \int_0^1 \int_0^x 1 \to k \int_0^1 [x]_0^x \to k \int_0^1 x \to k [x^2/2]_0^1 \to \frac{k}{2} = 1 \Rightarrow k = 2$
  - $f_x(x) = \int_0^x 2\,dy \to 2[y]_0^x \to 2x$ • $f_y(y) = \int_0^1 2\,dx \to 2[x]_0^1 \to 2$
  - Distribuzione X condizionata da $Y=y$: $f_{x|y}(x|y) = f(x,y)/f_y(y) \to 2/2 \to 1$
  - $f_x(x) \cdot f_y(y) = f_{xy}(x,y) \to 2x \cdot 2 \neq 2$ DIPENDENTI

- $f_{xy}(x,y) = \frac{1}{8}(6-x-y) \qquad 0<x<2 \; ; \; 2<y<4$
  - $f_x(x) = \int_2^4 f_{xy}(x,y)\,dy \to \int_2^4 6/8 + \frac{x}{8} - \frac{y}{8} \to (6-2x)/8$
  - $f_y(y) = \int_0^2 f_{xy}(x,y)\,dx \to \int_0^2 6/8 - \frac{x}{8} - y/8 \to (10-2y)/8$
  - Distribuzione X condizionata $Y=y$: $f_{x|y}(x|y) = f(x,y)/f_y(y) \to 1/8(6-x-y)/\frac{1}{8}(10-2y) \to (6-x-y)/(10-2y)$

- $f_{xy}(x,y) = k(x+y) \qquad 0<x<2 \; ; \; 0<y<2$
  - [VALORE K] $\int_0^2 \int_0^2 f_{xy}(x,y)\,dx\,dy = 1 \to k \int_0^2 \int_0^2 (x+y) \to k \int_0^2 [x^2/2 + xy]_0^2 \to k[2+2y]_0^2 \to 8k = 1 \to k = 1/8$
  - $P(X>Y) \to \int_0^2 \int_0^2 f_{xy}(x,y) \to \int_0^2 \int_0^2 \frac{1}{8}(x+y) \to 1/2$
  - $f_x(x) \to \int_0^2 f_{xy}(x,y)\,dy \to 1/4\,(x+1)$ • $f_y(y) \to \int_0^2 f_{xy}(x,y)\,dx \to 1/4\,(y+1)$
  - $f_x(x) \cdot f_y(y) = f_{xy}(x,y) \to \frac{1}{8}(x+y) \neq (1/4(x+1)) \cdot (1/4(y+1))$ DIPENDENTI

- $f_{xy}(x,y) = 1/8\,(x+y) \qquad f_x(x) = 1/4(x+1) \qquad f_y(y) = 1/4(y+1)$
  - $f_{x|y}(x|Y=y) \to f_{xy}(x,y)/f_y(y) \to 1/8(x+y)/\frac{1}{4}(y+1) \to 1/2\,(x+y/y+1)$
  - $F_{x|y}(x|Y=y) \to \int_0^x f_{x|y}(x|y)\,dz \to 1/2(y+1) \int_0^x z+y \to 1/2(y+1)\,[\frac{z^2}{2} + zy] \to x^2 + 2xy / 4(1+y)$

- $f_{xy}(x,y) = 2e^{-(x+y)} \qquad 0 \le x \le y \le \infty \iff y \ge x \ge 0 \; ; \; \infty \ge y \ge 0 \qquad Z = XY$
  - $f_x(x) = \int_0^\infty f_{xy}(x,y)\,dy \to [-2e^{-x-y}]_0^\infty \to 0 + 2e^{-x} \to 2e^{-x}$
  - $f_y(y) = \int_0^y f_{xy}(x,y)\,dx \to [-2e^{-x-y}]_0^y \to -2e^{-2y} + 2e^{-y}$
  - $f_x(x) \cdot f_y(y) = f_{xy}(x,y) \to 2e^{-x} \cdot (-2e^{-2y} + 2e^{-y}) \neq 2e^{-x-y}$ DIPENDENTE
  - $\to E(XY) = \int_0^\infty \int_0^y xy\, f_{xy}(x,y)\,dx\,dy \to 1/2$

- DENSITÀ di $S = X+Y \qquad X \sim P(\lambda x) \qquad Y \sim P(\lambda y) \qquad S \sim P(\lambda x + \lambda y)$?
  - $P(X+Y=S) \to \sum_{z=0}^{S} P(X=z)\,P(Y=S-z) \to \sum_{z=0}^{S} \frac{\lambda x^z}{z!} \cdot e^{-\lambda x} \cdot \frac{\lambda y^{S-z}}{(S-z)!} \cdot e^{-\lambda y}$

$$\to \frac{e^{(\lambda x + \lambda y)}}{S!} \cdot \sum_{z=0}^{S} \binom{S}{z} (\lambda x)^z \cdot (\lambda y)^{(z-S)} \to \frac{(\lambda x + \lambda y)^S}{S!} e^{-(\lambda x + \lambda y)}$$

- $X \sim \mathcal{P}(\lambda_X) \qquad Y \sim \mathcal{P}(\lambda_Y) \qquad Z = X+Y \qquad P(X|Z)\,? \qquad E(X|Z)\,?$

  - $P(X=x|Z=z) \to P(X=x|Z=z) \to P(X+Y=z|X=x)\,P(X=x)/P(X+Y=z) \to P(Y=z-x)\,P(X=x)/P(X+Y=z)$

    $\to \left[\frac{\lambda_Y^{z-x}}{(z-x)!}\cdot e^{-\lambda_Y}\right]\cdot\left[\frac{\lambda_X^{x}}{x!}\cdot e^{-\lambda_X}\right] \Big/ \left[\frac{(\lambda_X+\lambda_Y)^{z}}{z!}\cdot e^{-(\lambda_X+\lambda_Y)}\right] \to \binom{z}{x}\left(\frac{\lambda_Y}{\lambda_X+\lambda_Y}\right)^{z-x}\left(\frac{\lambda_X}{\lambda_X+\lambda_Y}\right)^{x} \to (X|Z=z) \sim Bi\left(z, \frac{\lambda_X}{\lambda_X+\lambda_Y}\right)$

  - $E(X|Z=z) = z\cdot\frac{\lambda_X}{\lambda_X+\lambda_Y}$

- $(X|Y=y) \sim Bi(y, 1/3) \qquad y = 1,2 \qquad x = 0,1,2,\dots$

  - $P(X=x|Y=y) = \binom{y}{x}\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{y-x} \to P(X=0|Y=1) \sim Bi(1,1/3) \to \begin{matrix}\binom{1}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^1 = 2/3 \\ \binom{1}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^0 = 1/3\end{matrix} \to X\begin{cases} 0 & 1 \\ 2/3 & 1/3\end{cases}$

  - $P(X=x|Y=y)\cdot P(Y=y)$ [PROBABILITÀ CONGIUNTA]

$x \backslash y$	1	2
0	2/12	1/3
1	1/12	1/3
2	0	1/12

$P(X=0|Y=1)P(Y=1) \to \binom{1}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^1\cdot\frac{1}{4} = 2/12$

$P(X=0|Y=2)P(Y=2) \to \binom{2}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^2\cdot 3/4 = 1/3$

$P(X=1|Y=1)P(Y=1) \to \binom{1}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^0\cdot 1/4 = 1/12$

$P(X=1|Y=2)P(Y=2) \to \binom{2}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^1\cdot 3/4 = 1/3$

$P(X=2|Y=1)P(Y=1) \to \binom{1}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^{1-2}\cdot 1/4 = 0$

$P(X=2|Y=2)P(Y=2) \to \binom{2}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^0\cdot 3/4 = 1/12$

  - $(Y|X=1) = \begin{cases} 1 & 2 \\ \frac{1/12}{5/12} & \frac{1/3}{5/12} \end{cases}$ ← marginale $X=1$ $\to \begin{cases} 1 & 2 \\ 1/5 & 4/5\end{cases}$

  - $E(Y|X=1) = 1\cdot\frac{1}{5} + 2\cdot 4/5 = 9/5$

  - $Pr(Y>X)$

- $F_{XY}(x,y) = [1-(x+1)e^{-x}][1-(y+1)e^{-y}] \qquad x>0;\ y>0$

  - [VERIFICA] $\lim_{\substack{x\to\infty \\ y\to\infty}} F_{XY}(x,y) = 1 \quad ; \quad \lim_{\substack{x\to 0 \\ y\to 0}} F_{XY}(x,y) = 0$

  - $F_X(x) = \lim_{y\to\infty} F_{XY}(x,y) \to 1-(x+1)e^{-x}$
  - $F_Y(y) = \lim_{x\to\infty} F_{XY}(x,y) \to 1-(y+1)e^{-y}$

  - $f_{XY}(x,y) = \Delta^2/\Delta x\Delta y\ F_{XY}(x,y) \to \frac{\Delta}{\Delta x}F_X(x)\cdot\frac{\Delta}{\Delta y}F_Y(y) \to (-e^{-x}+(x+1)\cdot e^{-x})(-e^{-y}+(y+1)\cdot e^{-y})$

    $\to xe^{-x}\cdot ye^{-y}$

  - $f_X(x) = \int_0^\infty f_{XY}(x,y)\,dy = xe^{-x}$
  - $f_Y(y) = \int_0^\infty f_{XY}(x,y)\,dx = xe^{-y}$

  - $f_X(x)\cdot f_Y(y) = f_{XY}(x,y) \to$ SONO INDIPENDENTI

  - $P(X>2Y) \to \int_0^\infty\int_{2y}^\infty f_{XY}(x,y)\,dx\,dy \to \int_0^\infty\int_{2y}^\infty f_X(x)\cdot f_Y(y)\,dx\,dy \to \int_0^\infty f_Y(y)\cdot[F_{(x)}(x,y)]_{2y}^{\infty}$

    $\to \int_0^\infty y\cdot e^{-2y}(1+2y)\,e^{-2y}\,dy \to 7/27$

- $X_1 \dots X_{80} \sim f(x) = 3x^2 \qquad x \in (0,1) \qquad P(S_{80} < 65) = ? \qquad n = 80$
  - $E(x) = \int_0^1 x \cdot 3x^2 \, dx \to 3/4$
  - $E(x^2) = \int_0^1 x^2 \cdot 3x^2 \to 3/5$
  - $VAR(x) = E(x^2) - E(x)^2 = 3/80$
  - $P(S_{80} \leq 65) \to P\left(\frac{S_{80} - (n \cdot E(x))}{\sqrt{n \cdot VAR(x)}} < \frac{x - (E(x) \cdot n)}{\sqrt{n \cdot VAR(x)}}\right) \to P(S_{80} < 2.887) \to \Phi(2.887) = 0.998$

- $X \sim N(11.4,\ 1.3) \qquad n = 20$
  - $P(9.7 \leq \bar{x} \leq 11) \to P\left(\sqrt{20} \cdot \frac{9.7-11.4}{1.3} < x < \sqrt{20} \cdot \frac{11-11.4}{1.3}\right) \to (\Phi(-1.377)) - (\Phi(-5.84))$

    $\to \cancel{1-}0.9147 - (8.04) \to 0.844$

  - $P(a \leq \bar{x} < b) = 0.8 \to P\left(\sqrt{20} \cdot \frac{(a-11.4)}{1.3} < Z < \sqrt{20} \cdot \frac{(b-11.4)}{1.3}\right) = 0.8 \to P\left(Z < \frac{\sqrt{20}}{1.3}(b-11.4)\right) + P\left(\frac{\sqrt{20}}{1.3}(a-11.4)\right) - 1$

    $\to 2P\left(Z < \frac{\sqrt{20}}{1.3}(b-11.4)\right) = 0.8 \to \Phi\left(\frac{\sqrt{20}}{1.3} \cdot (b-11.4)\right) = \frac{0.8+1}{2}$

    $\to \frac{\sqrt{20}}{1.3}(b-11.4) = 1.28 \to (b-11.4) = 1.28 \cdot \frac{1.3}{\sqrt{20}} \to 0.372$

    $\to P(11.4 - 0.372 < \bar{x} < 11.4 + 0.372) = 0.8 \to P(11,0275 < \bar{x} < 11,7725) = 0.8$

- $X \sim N(175,\ 4\overset{\sigma^2}{2}) \qquad n = 50 \qquad$ chebychev: $P(|x - \overset{E(x)}{\mu}| \geq \varepsilon) \leq \frac{\overset{\sigma^2/n}{VAR(x)}}{\varepsilon^2}$
  - $P(170 < \bar{x} < 180) \to P(-5 < \bar{x} - 175 < 5) \to P(|\bar{x} - 175| < 5) \to 1 - P(|\bar{x} - 175| > 5)$

    $\to P(|\bar{x} - 175| > \overset{\varepsilon}{5}) \leq VAR(x)/\overset{\varepsilon^2}{25}$ [CHEBYCHEV]

    $\to 1 - P(|\bar{x} - \mu| > 5) \geq 1 - (42/50)/25 \to P(|\bar{x} - \mu| > 5) \geq 1 - 0.0336 \to 0.9664$

  - [n=?] $P(174 < \bar{x}_n < 176) \geq 0.9 \to P(-1 < \bar{x}_n - \mu < 1) \geq 0.9 \to P(|\bar{x}_n - \mu| < 1) \geq 0.9$

    $\to 1 - P(|\bar{x}_n - \mu| > 1) \geq \cancel{1} - \frac{42/n}{1}$

    $\to 1 - \frac{42}{n} \geq 0.9 \to n \geq 420$

COMANDI DI R

ARITMETICA

LS( ) = controllare cosa c'è nella directory

RM (nome) = rimuovere oggetto "nome"

RM (nome1, nome2, ...) = rimuovere più oggetti

RM (LIST = LS( )) = rimuovere tutto quello fatto fin'ora quando si inizia un nuovo progetto

+, −, /, * = quattro operazioni

**, ^ = elevamento a potenza

EXP = esponenziale

SQRT (numero) = radice quadrata

LOG = logaritmo naturale

SIN( ), COS( ) = seno e coseno

ASIN( ), ACOS( ) = forme inverse

ABS = valore assoluto

ASSEGNAZIONE VALORI

X ← 5 = salvo in x il valore 5

log (x) → y = assegno ad y il valore di log (5)

Z = x + y = in Z avremo il valore della somma di x + y (5 + log (5))

VALORI LOGICI

X ← 10 , X > 10 = assegno 10 ad x, e verifico se X > 10 (FALSE)

& = intersezione (TT = T, TF = F, FT = F, FF = F)

| = unione (TT = T, TF = T, FT = T, FF = F)

!= = diverso

!X = inverso di X (T = F, F = T)

CREAZIONI VETTORI

X ← c(2, 3, 5, 7, 11) = si crea in x il vettore c contenente (2, 3, 5, 7, 11)

X ← scan( ) = quando si hanno tanti vettori, ti chiede di inserirne finchè vuoi

X ← scan ("nomefile.dat") = puoi importare un vettore da file e lo assegno ad x

SUCCESSIONI

XX ← 1 : 10 = crea vettore con la sequenza di numeri da 1 a 10

XX ← 100 : 1 = crea vettore con sequenza di numeri in ordine decrescente da 100 a 1

XX ← seq (FROM = 100, TO = 1) = altro metodo per sequenza da 100 a 1 in ordine decrescente

SEQ (0,1 , by = 0.1) = creo sequenza che va da 0 a 1 crescendo di 0,1 in 0,1 [0.0, 0.1, 0.2, ..., 1.0]

rep (2 , TIMES = 3) = vettore con 2 ripetuto tre volte [2 2 2] ≡ rep (2, 3)

x ← c (rep(2,3), 4, 5, rep(1,5)) = [2 2 2 4 5 1 1 1 1 1]

x ← 1:10, x*2 , x > 5 = [2 4 6 8 10 12 14 16 18 20] [F F F F F T T T T T]

ESTRAZIONI ELEMENTI da un VETTORE

xx ← 100:1 , xx[7] = estrae il settimo valore della serie [94]

xx ← 100:1 , xx[c(2,3,5,7,11)] = estrae il secondo, terzo, ..., undicesimo elemento della serie [99 98 96 94 90]

xx ← 100:1 , xx [91:85] = estrae dal novantunesimo all'ottantacinquesimo elemento [10 11 12 13 14 15 16]

yy ← xx[c(2,3,5,7,11)] = assegno ad yy i valori estratti da xx

xx ← c(1,2,4,8,16,32) , x[-4] = se si richiede un numero negativo, quella posizione verrà omessa [1 2 4 16 32]

x ← 1:10 , length(x) = stampo la lunghezza del vettore [10]

x ← 1:10 , MAX(x) , MIN(x) = stampo il massimo [10] e il minimo [1] del vettore

SUM(x) = sommatorio PROD(x) = produttorio

SORT(x) = ordino la lista in ordine crescente

INDICATORI CATEGORIA

esperimento ← FACTOR (c('a', 'b', 'b', 'c', 'a', 'b')) = assegno ad esperimento il nome per ogni prova

LEVELS (esperimento) = stampo il nome delle prove ["a" "b" "c"]

risposta ← c(10, 3, 7, 6, 4, 5) = assegno alle prove il valore dell'esperimento

risposta [esperimento == 'a'] = stampo i risultati di una specifica prova [10 4]

MATRICI

x ← MATRIX (c(2,3,5,7,11,13), ncol = 2) = creo una matrice con 2 colonne e per default 3 righe $\begin{bmatrix} 2 & 7 \\ 3 & 11 \\ 5 & 13 \end{bmatrix}$

X2 ← scan ('nomematrix1', sep = ',') = carico i dati per una matrice da un file nomematrix1

mx ← MATRIX (X2, ncol = 5 , byrow = T) = creo la matrice con i dati caricati, con numero colonne 5 e numero righe in default in base al numero di elementi

x [2, 1] = estrae elemento della matrice [nrow, ncol]

x [, 1] = si ottiene l'intera colonna 1

x [c(1,4), c(3,4)] = verrà ricreata una matrice (prendendo valori da x) con le coordinate (1,3), (1,4) (4,3), (4,4)

dim (x) = restituisce la dimensione della matrice x (nrow, ncol)

ELEMENTI di PROGRAMMAZIONE

nomefunzione <- function (a, b, c, ...)   STRUTTURA BASE di DICHIARAZIONE FUNZIONE

- cubo <- function (x) {
  return(x^3)
  }

  cubo(3)

  [27]

- MEDIA <- function (x) {
  y <- 0
  for (i in 1: length(x)) {
  y <- y+x [i] }
  y <- y / length(x)
  return (y) }

  ≡ MEDIA <- function (x) {
  sum (x) / length(x)
  }

  ≡ mean(x)

STATISTICA DESCRITTIVA

TABELLE di FREQUENZA

- maschi <- scan ('maschi.dat') = importo dati
- table (maschi) = creo la tabella

165	166	179	...	183	185	...	153
1	2	5		3	6		1

- classi <- 150+5*(0:10) = creo le classi   150 155 160 ... 200

cut (maschi, breaks = classi) = assegno maschi ad ogni classe   (185, 190] (180, 185] ... (170, 175] ... [175, 180] ... (185, 200]

Table (cut (maschi, breaks=classi) = creo tabella di frequenza

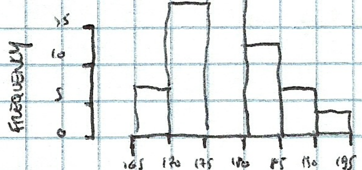
(150, 155]	...	(165, 170]
0		7

Table (cut (maschi, breaks=10)) = crea tabella di frequenza con classi default

freqcum <- cumsum (table (cut (maschi, breaks = classi)))/length (maschi) = frequenze cumulative relative

GRAFICI

hist (maschi) = creo un istogramma basandomi sui dati della tabella maschi



hist (maschi, breaks = classi)
hist (maschi, breaks = 10)
} creo istogrammi in base alle classi frequenze sempre 0:15

INDICI di POSIZIONE e VARIABILITÀ

- min = fornisce il valore della più piccola osservazione campionaria
- max = fornisce il valore della più grande osservazione campionaria
- range = fornisce gli estremi del campo campionario (max e min)
- mean = calcola la media
- median = calcola la mediana
- var = calcola la varianza campionaria
- quantile = calcola il quantile
- summary = fornisce una tabella con resoconto delle voci sopracitate

quantile (maschi) =

0%	25%	50%	75%	100%
165	174	178	183	193

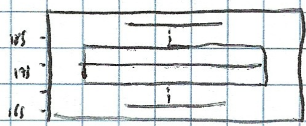
quantile (maschi, probs = c(.15, .47)) = avremo solo il quantile 15% e 47%

15%	47%
172	178

summary (maschi) =

MIN	1ST QUA	MEDIAN	MEAN	3RD QUART	MAX
165.000	174.000	178.000	178.1538	183.00	193.000

BOXPLOT (maschi) = diagramma con le informazioni di summary [diagramma a scatola]



PROBABILITÀ

CALCOLO COMBINATORIO

- $n!$ ⇒ prod (1:5) ⇒ 6! ⇒ [120]
- $6!/3$ ⇒ $D_{6,3}$ ⇒ $6 \times 5 \times 4$ ⇒ prod (6:(6-3+1)) ⇒ [120]
- $\binom{n}{k}$ ⇒ choose (4,2) ⇒ 4x3 / 2x1 ⇒ [6] coefficiente binomiale

DISTRIBUZIONI di PROBABILITÀ

nome → abbreviazione → parametri → default

- NORMALE ⇒ norm ⇒ mean, sd ⇒ 0, 1
- GAMMA ⇒ gamma ⇒ shape ⇒ –
- UNIFORME ⇒ unif ⇒ min, max ⇒ 0, 1
- BINOMIALE ⇒ binom ⇒ n, p ⇒ – –
- POISSON ⇒ pois ⇒ λ ⇒ –
- ESPONENZIALE ⇒ exp ⇒

$X \sim Bin(10, 0.2)$ $x = 2$

- dbinom (2, 10, 0.2) = probabilità che X assuma il valore x = 2 [DENSITÀ] [0.3019899]
- pbinom (2, 10, 0.2) = funzione di ripartizione $P(X \leq x)$ [0.6777995]
- pbinom (2, 10, 0.2, lower.tail = F) = funzione ripartizione $P(X > x)$ [0.3222005]
- qbinom (0.45, 10, 0.2) = otteniamo il quantile 0.45 (45%) [1]

SIMULAZIONE VAR CASUALI

CAMPIONAMENTO

- sample (1:6, 10, replace = T) = sample permette di estrarre (con o senza reinserimento) un certo numero di valori da un insieme prefissato. il nostro esempio simula 10 lanci di un dado (con doppioni)

- urna <- c('b1', 'b2', 'b3', 'b4', 'n1', 'n2', 'n3') = creiamo un'urna (4 palline bianche e 3 nere)

  sample (urna, 2) = estraiamo 2 palline dall'urna senza reinserimento

  sample (urna, 2, replace = T) = estraiamo 2 palline dall'urna con reinserimento

GENERARE NUMERI PSEUDO-CASUALI

$y_{n+1} = (a y_n + b) \bmod m$ ← resto divisione per m

$x_0 = 83$ $a = 1573$ $b = 19$ $m = 10^3$

$x_1 = 140016$ $(\bmod 1000) = 16$

$x_2 = 25187$ $(\bmod 1000) = 187$

⋮

≡

in R:

```
n <- 5
y <- numeric(n+1)        # seme
y[1] <- 83
a <- 1573
~~mk~~ <- 10^3
b <- 19
for (i in 2:(n+1))
  y[i] <- (a * y[i-1] + b) %% m
y <- y/m
y
```

set.seed ( ) = genera il seme
runif ( )

- ~~x <-~~ rnorm (1000) = simulare 1000 valori di una distribuzione normale standard

  plot (density (x)) = produce uno stimo grafico del vettore x

- par (mfrow = c(1,2))
  plot (density (x))
  curve (dnorm (x), xlim = c(-3.5, 3.5))
  par (mfrow = c(1,1))

  confronto fra la densità della normale simulata e quella di una vera normale

- a <- 1.5
  b <- 1.5
  n <- 1000
  set.seed (567)
  y <- rbeta (n, a, b)
  hist (y, prob = True)
  curve (dbeta (x, a, b), from = 0, to = 1, ylab = 'f(y)', xlab = 'y', add = True)

  Simulazione di 1000 valori indipendenti da una variabile Beta

  $f(y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1} (1-y)^{b-1}$ $0 < y < 1$

  $a, b = 1.5$

TEOREMA LIMITE e APPLICAZIONI

LEGGE GRANDI NUMERI

$X_i \quad i = 1, \dots \qquad E(X_i) = \mu \qquad \bar{X}_m = \left(\sum_{i=1}^{m} X_i\right)/m \rightarrow$ converge molto probabilmente ad μ

media campionaria

$P_r\left(\lim_{m \to \infty} X_m = c\right) = 1 \leftarrow$ convergenza quasi certa di una successione X_i ad una costante c

$m = 10 \qquad Y \sim \mathcal{P}(5)$

```
set.seed(30)
y <- rpois(10, 5)
mean(y)
```

] si fissa la variabile casuale da usare
si generano m valori casuali, si calcola la media e si itera il procedimento aumentando m di volta in volta

```
y <- c(y, rpois(10,5))
mean(y)
```

] raddoppio

```
y <- c(y, rpois(20,5))
mean(y)
```

] raddoppio ancora

```
mobs <- (1 : length(y))
m <- cumsum(y) / mobs
plot(mobs, m, type = 'l')
```

] grafico andamento $\bar{X}_m$

METODO MONTECARLO

$\left[I(f) = \int_0^1 f(y)\,dy\right]$

$\rightarrow$ con legge grandi numeri, calcolo in modo approssimativo certe quantità

$$I(f) = \int_0^1 \frac{1}{\sqrt{2\pi}} \cdot e^{-y^2/2}\,dy$$

[funzione ripartizione normale standard]

```
set.seed(1)
m <- 1000
y <- runif(m)
a <- mean(exp(-y^2/2))/sqrt(2*pi)
a
```

[0.341437]

$\equiv$ pnorm(1) - pnorm(0)

[0.341437]

TEOREMA LIMITE CENTRALE

$X_i \quad i = 1, \dots$ successioni variabili casuali iid con media μ e varianza σ^2

$[Z = (\bar{X}_m - \mu)/(\sigma/\sqrt{m})]$

$X_i \sim Bi(1, 0.2) \qquad \sigma^2 = VAR(X_i) = p(1-p) = 0.16$

All'aumentare di m

$\bar{Z}_m = (\bar{X}_m - \mu)/(\sigma/\sqrt{m})$ converge in distribuzione $N(0,1)$

```
for (mfrow = c(2,2))
  for (m in mobs)
{   y <- 0:m
    prob <- pbinom(y, m, p)
  z <- (y/m-p)*sqrt(m)/sqrt(p*(1-p))
  ind <- (z>-3) & (z<3)
    z <- z[ind]
    prob <- c(0, prob[ind])
  plot(stepfun(z, prob, f = 0), verticals = FALSE, pch = 20, main=paste("m= ", m, ", p= ", p), ylab = "F(z)", xlab = "z")
```

curve(pnorm(x), from = min(z), to = max(z), add = TRUE) }