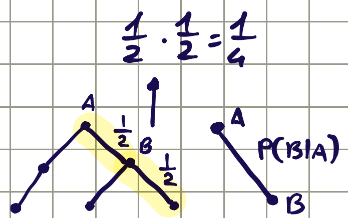
STATISTICA DESCRITTIVA

- $var(x)$: $\frac{1}{n-1} \cdot \sum (x_i - \bar{x})^2$
- $var_e(x)$: $\frac{1}{n} \cdot \sum (x_i - \bar{x})^2$, se si conosce f. relativa $\sum x_i^2 \cdot \frac{f_i}{n} - \bar{x}^2$
- rapporto tra $var(x)$ e $var_e(x)$ $\Longrightarrow$ $var(x) = \frac{n}{n-1}(var_e(x))$
- $\sigma(x)$ dev. standard: $\sqrt{var(x)}$
- $\sigma_e(x)$ dev. standard emp.: $\sqrt{var_e(x)}$
- e.c.d.f.: $F_e(t) = \frac{\# \{i \mid x_i \leq t\}}{n}$
- $cov(x,y)$: $\sum \frac{(x_i - \bar{x})(y_i - \bar{y})}{n-1}$ OSS: $\frac{x_i = y_i \ \forall i}{cov(x,y) = var(x) = var(y)}$
- $cov_e(x,y)$: $\sum \frac{(x_i - \bar{x})(y_i - \bar{y})}{n}$
- relazione tra $cov(x,y)$ e $cov_e(x,y)$: $cov(x,y) = \frac{n}{n-1}(cov_e(x,y))$
- $r(x,y)$ coeff. correl.: $\frac{cov(x,y)}{\sigma(x) \cdot \sigma(y)}$ deb. ≈ 0, forte $\approx 1/-1$, perf. $= 1/-1$
- retta di regress.: $y = m^* x + q^*$ $\begin{cases} m^* = \frac{cov(x,y)}{var(x)} \\ q^* = \bar{y} - m^* \cdot \bar{x} \end{cases}$
- percentile: $\frac{K}{100} \cdot n$

PROBABILITÀ E INDIPENDENZA

- $P(A)$ Laplace: $\frac{\# A}{\# \Omega}$
- $P(A^c)$: $1 - P(A)$
- $P(A \cup B)$: $P(A) + P(B) - P(A \cap B)$
- $P(A \mid B)$: $P(A \cap B) / P(B)$
- $P(A \cap B)$: $P(A \mid B) \cdot P(B)$
- $B \subseteq A \Longrightarrow P(A \setminus B) = P(A) - P(B)$
- Formula di fattoriz.: $P(A) = \sum P(A \mid B_i) \cdot P(B_i)$ $\longrightarrow$



- Formula di Bayes: $P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$
- Definizioni indipendenza: $P(A \cap B) = P(A) \cdot P(B)$

CALCOLO COMBINATORIO

- SI ordine, NO rimpiazzo: $\frac{n!}{(n-k)!}$
- SI ordine, SI rimpiazzo: n^k
- NO ordine, NO rimpiazzo: $\binom{n}{k} \Rightarrow \frac{n!}{k!(n-k)!}$
- NO ordine, SI rimpiazzo: $\binom{n-1+k}{k} \Longrightarrow \frac{(n-1+k)!}{k!\,(n-1)!}$

- esattamente k elementi su n: $P(A) = \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{k}}$

VARIABILI ALEATORIE

- X, Y equidistribuite: $P_X = P_Y$ (leggi di distribuzione)
- Funzione di massa: $P_X : \{a_1, \dots, a_n\} \to \mathbb{R}$
  - discreta: $\sum P_X(a_i)$
    - oss: $P_X(a_i) \geq 0 \ \forall i$, $\sum P_X(a_i) = 1$
  - continua: $A = [a,b]$ $P(X \in A) = \int_a^b f(x)dx$
    - oss: $f(x) \geq 0 \ \forall x$, misura area sotto grafico $f = 1$

- Funzione di ripartizione (c.d.f): Proprietà $F(x)$:
  - (x) = devono essere continue
  - proprietà:
    - $X_1 \leq X_2 \Longrightarrow F(X_1) \leq F(X_2)$
    - $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$
    - $X_n \to X \Rightarrow F(X_n) \to F(x)$ (è continua a destra)
  - continua: $\int_{-\infty}^{x} f(x)dx$
  - discreta: funzione a salti con i salti corrispondenti ai punti in cui la funzione di massa non si annulla

Funzione di massa: assegna una probabilità a ciascun valore che una v.a. può assumere

Funzione di riport.: dato un valore x, restituisce la prob. che una v.a. sia $\leq x$

- $B(1,p)$ Bernoulli: esperimento dicotomico
  - $P_x(1) = p$, $P_x(0) = 1-p$
  - $E[x] = \sum a^k \cdot p(x=a_k) = p$
  - $E[x^n] = \sum a_k^n \cdot p(x=a_k) = p$
  - $var(x) = p(1-p)$
- $B(n,p)$ Binomiale: n prove con risult. dicotomico
  - $P(x=k) = \binom{n}{k} p^k (1-p)^{n-k}$
  - $E[x] = \sum_i E[X_i] = np \implies E[x^k] = \sum_x x^k P(X=x)$
  - $var(x) = np(1-p)$
- $G(p)$ Geometrica: successo nell'istante T
  - $P_T(k) = p(1-p)^{k-1}$
  - assenza memoria: $P(T=n+k \mid T>n) = P(T=k)$
  - $E[x] = \frac{1}{p}$
  - $var(x) = \frac{1-p}{p^2}$
  - Se $X \geq 0 \implies \exists\, E[x]$
- $Poiss(\lambda)$: conta il # successi, λ è il # medio di eventi nell'intervallo
  - $P_x(k) = P\{x=k\} = \frac{\lambda^k \cdot e^{-\lambda}}{k!}$
  - $E[x] = \lambda$
  - $var(x) = \lambda$
- $IP(N,F,n)$ Ipergeometrica: # successi, No reinserimento
  - N unità, F favorevoli, n unità estratte
    - $P_x = \frac{\binom{F}{k}\binom{N-F}{n-k}}{\binom{N}{n}}$
  - $E[x] = n \cdot \frac{F}{N}$
  - $var(x) = n \cdot \frac{F}{n}\left(1 - \frac{F}{N}\right)\frac{N-n}{N-1}$

- $U(a,\beta)$ Uniformi: valori costanti in un intervallo
  - f densità $\mathbb{1}(a,\beta)\begin{cases}\frac{1}{\beta-a} & \text{per } x\in(a,\beta)\\ 0 & \text{altrimenti}\end{cases}$
  - $P_X = \dfrac{\text{length}((a,\beta)\cap(a,b))}{\beta-a}$
  - $E[x]=(\beta+a)/2$
  - $E[x^2]=\frac{1}{3}(a^2+a\beta+\beta^2)$
  - $var(x)=\dfrac{(\beta-a)^2}{12}$
- $Exp(\lambda)$ Esponenziali: tempi di attesa evento, $\lambda>0$
  - f densità $\begin{cases}\lambda e^{-\lambda x} & \text{per } x>0\\ 0 & \text{per } x\le 0\end{cases} \Rightarrow f(x)=\lambda e^{-\lambda x}\cdot \mathbb{1}_{(0,\infty)}(x)$
  - $E[x]=1/\lambda$

    $\hookrightarrow P\{x\le k\}=\int_{-\infty}^{k} f(x)$ con $-\infty$ da adattare al range effettivo. Se min=0, $\int_0^1 f(x)$
  - $E[x^2]=2/\lambda^2$
  - $var(x)=1/\lambda^2$
  - densità: $1-e^{-\lambda x}$
- $N(m,\sigma^2)$ Gaussiana: probabilità più alta intorno alla media
  - $\hookrightarrow$ $\underline{Z\sim N(0,1)}$: caso standard
    - $P_X=1$
    - $E[x]=0$
    - $E[x^2]=1$
    - $var(x)=1$
  - $\hookrightarrow$ $\underline{N(m,\sigma^2)}$: ci si riconduce a $N(0,1)$ tramite standardizzazione
    - standard.: $Z=\dfrac{X-m}{\sigma}$ da cui si può calcolare $\Phi(z)$
    - $E[x]=m$
    - $Var(x)=\sigma^2\underbrace{var(z)}_{=N(0,1)}=\sigma^2$

STATISTICA DESCRITTIVA

- $E[X] = \mu \begin{cases} \sum_i a_i \cdot P(a_i) < \infty & \text{discrete} \\ \int_{-\infty}^{+\infty} x \cdot f(x)\,dx = \int_{-\infty}^{+\infty} x\, \mathbb{1}_{(\alpha,\beta)}(x)\,dx < \infty & \text{continua} \end{cases}$
  - Proprietà: X e Y v.a. con $E[X]$, a e $b \in \mathbb{R}$
    - $X+Y$ ha $E[X+Y] = E[X] + E[Y]$
    - $aX+b$ ha $E[aX+b] = aE[X]+b$
    - $X \geq 0 \Longrightarrow E[X] \geq 0$, $X \geq Y \Longrightarrow E[X] \geq E[Y]$
- $E[X^n]$ momento di ordine n di X:
  - caso discreto: $E[X^n] = \sum_k a_k^n \cdot p(X = a_k)$
  - caso continuo: $E[X^n] = \int_{-\infty}^{+\infty} x^n f(x)\,dx$
- Proprietà $Var(X)$
  - $Var(X) = 0 \Longleftrightarrow X$ costante
  - $Var(X) = E[X^2] - E[X]^2$
  - $Var(aX+b) = a^2 Var(X)$
- Disuguaglianza di Markov: X v.a. ≥ 0, $a > 0 \Longrightarrow P\{X \geq a\} \leq \frac{1}{a} E[X]$
- Disuguaglianza di Chebyshev: $P\{|X - E[X]| \geq d\} \leq \frac{1}{d^2} Var(X)$
- $E[XY] = E[g(X,Y)] = \sum_i \sum_j g(x_i, y_j) \cdot P(X = x_i, Y = y_j)$
  - se sono indipendenti si può semplificare e si ha:
    $E[XY] = E[X] \cdot E[Y]$

VARIABILI ALEATORIE DOPPIE

- $P\{(X,Y)\in A\} = \sum_{(x_i,y_i)\in A} P(x_i, y_j)$
  - proprietà:
    - $P(x_i, y_j) \geq 0 \quad \forall i,j$
    - $\sum_i \sum_j P(x_i, y_j) = 1$
- X, Y discrete, indipendenti $\iff P_{(X,Y)}(x_i, y_j) = P_X(x_i)\cdot P_Y(y_j) \quad \forall (x_i, y_j)$
- $X \sim B(m,p)$ e $Y \sim B(h,p)$ indipendenti $\Rightarrow X+Y \sim B(m+h, p)$
- $X \sim P(\lambda)$ e $Y \sim P(\mu)$ indipendenti $\Rightarrow X+Y \sim P(\lambda+\mu)$
- $X \sim N(m_1, \sigma_1^2)$ e $Y \sim N(m_2, \sigma_2^2)$ indip. $\Rightarrow X+Y \sim N(m_1+m_2, \sigma_1^2+\sigma_2^2)$
- X, Y v.a. con $E[X]$ e indip. $\Rightarrow X\cdot Y$ ammette $E[X\cdot Y] = E[X]\cdot E[Y]$

V.A CON MOMENTO SECONDO

Date X, Y, Z v.a. con momento secondo:

- $Cov(X,Y)$: $E[(X-E[X])(Y-E[Y])] = E[XY] - E[X]E[Y]$
- ρ coeff. corr.: $\dfrac{cov(x,y)}{\sqrt{var(x)\,var(y)}}$
- Scorrelate $\Leftarrow \rho(x,y) = 0$ oppure se X e Y
- Indipendenti $\overset{\Rightarrow}{\not\Leftarrow}$ Scorrelate
- X, Y scorrelate $\iff var(x+y) = var(x) + var(y)$
- Bilineare: $cov(aX + bY + c, Z) = a\cdot cov(X,Z) + b\cdot cov(Y,Z)$
- Simmetria:
  - $cov(X,Y) = cov(Y,X)$
  - $cov(X,X) = var(X)$
- Convergenza: $\forall \varepsilon > 0 \quad \lim_{i\to\infty} P\{|X_i - X| > \varepsilon\} = 0$

STIMATORI MLE

- $L(\lambda; x_1, \dots, x_n) = P_\lambda(x_1) \cdot \ldots \cdot P_\lambda(x_n)$
- densità: Si calcola facendo la derivata della c.d.f
  - $f(x) \geq 0$
  - $\int_{-\infty}^{+\infty} f(x)\,dx = 1$
- stimatore corretto: $E[\hat{\theta}] = \theta$
- stimatore consistente: $\hat{\theta}_n \xrightarrow[n \to \infty]{P} \theta$

- stimatori noti:

  $E[Var(x)] = E[\delta^2] = E[\sum^n (x_i - \bar{x})^2] = (n-1)\delta^2$