FRAME 0.1. Contents

1. Analisi combinatoria

2. Assiomi della probabilità

3. Probabilità condizionata ed indipendenza

4. Variabili aleatorie discrete

5. Variabili aleatorie continue

6. Leggi congiunte

6.1. Distribuzioni condizionate.

FRAME 6.1. *Siano X ed Y variabili discrete di densità p_X e p_Y e densità congiunta p. Condizioniamo X all'evento $Y = y$: poniamo*

$$p_{X|Y}(x \mid y) \stackrel{def}{=} P\{X = x \mid Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}} =$$

$$= \frac{p(x,y)}{p_Y(y)}.$$

In maniera analoga si definisce la distribuzione condizionata

$$F_{X|Y}(x \mid y) = P\{X \leq x \mid Y = y\} = \sum_{a \leq x} p_{X|Y}(a \mid y).$$

FRAME 6.2. ***Esempio*** *Siano X ed Y variabili discrete con densità congiunta $p(x,y)$ definita da*

$$p(0,0) = 0.4,\ p(0,1) = 0.2,\ p(1,0) = 0.1,\ p(1,1) = 0.3.$$

Determinare la densità discreta di X e la densità condizionata di X, dato $Y = 1$.

FRAME 6.3. *Siano X ed Y variabili continue di densità f_X ed f_Y, e densità congiunta f. Supposto $f_Y(y) > 0$, condizioniamo X all'evento $Y = y$: poniamo*

$$f_{X|Y}(x \mid y) \stackrel{def}{=} \frac{f(x,y)}{f_Y(y)}.$$

Allora

$$P\{X \in A \mid Y = y\} = \int_A f_{X|Y}(x \mid y)dx.$$

In particolare

$$F_{X|Y}(x \mid y) = P\{X \leq x \mid Y = y\} = \int_{-\infty}^{x} f_{X|Y}(x \mid y)dx.$$

FRAME 6.4. ***Esempio****. Siano X ed Y variabili continue con densità congiunta*

$$f(x,y) = \begin{cases} \frac{5}{12}x(2-x-y) & \text{se } 0 < x < 1,\ 0 < y < 1, \\ 0 & \text{altrimenti.} \end{cases}$$

Si calcolino la densità di X e la densità condizionata di X dato $Y = 1/4$.

FRAME 6.5. ***Un esempio*** *Una macchina produce monete. La taratura della macchina è difettosa: pertanto produce monete con baricentro variabile. Prendiamo una moneta prodotta dalla nostra macchina e lanciamola ripetutamente m volte. La probabilità X che in un singolo lancio esca testa è costante nelle m prove, ma dipende dalla posizione del baricentro, ed è pertanto una variabile aleatoria continua: denotiamo con f_X la sua densità. Sia N la variabile discreta che conta il numero di successi nelle m prove.*

FRAME 6.6. *Non è chiaro cosa sia la densità di N; è invece chiaro che*

$$P\{N = n \mid X = x\} = \binom{m}{n} x^n (1-x)^{m-n}.$$

Diciamo che $p_{N|X}(n \mid x) = P\{N = n \mid X = x\}$ è la densità di N condizionata ad X.

FRAME 6.7. *Il nostro esperimento si suddivide in due fasi indipendenti: nella prima si determina la probabilità di successo x, nella seconda si contano i successi. Applicando il principio di moltiplicazione possiamo definire la densità congiunta di X ed N:*

$$h_{X,N}(x,n) = \lim_{\Delta \to 0} \frac{1}{\Delta} P\{x \leq X \leq x + \Delta, N = n\} =$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta} P\{N = n \mid x \leq X \leq x + \Delta\} P\{x \leq X \leq x + \Delta\} =$$

$$= P\{N = n \mid X = x\} \lim_{\Delta \to 0} \frac{1}{\Delta} P\{x \le X \le x + \Delta\} = p_{N|X}(n \mid x) f_X(x).$$

FRAME 6.8. *Dalla densità congiunta ricaviamo facilmente le densità marginali:*

$$f_X(x) = \sum_n h_{X,N}(x,n) = \sum_n p_{N|X}(n \mid x) f_X(x)$$

e, di maggior interesse,

$$p_N(n) = \int_{-\infty}^{\infty} h_{X,N}(x,n)dx = \int_{-\infty}^{\infty} p_{N|X}(n \mid x) f_X(x)dx.$$

FRAME 6.9. *Voglio andare a votare per i 4 referendum abrogativi sulla legge per la fecondazione medicalmente assistita partendo da casa all'ora N; sono infatti incerto se partire alle 14, alle 15 o alle 16. Supponiamo che il tempo in minuti impiegato per recarsi al seggio e votare per i 4 quesiti sia una variabile normale di parametri $(N, 1)$. A che ora X avrò votato?*

FRAME 6.10. *Chiaramente X è una variabile aleatoria continua. Non è chiaro cosa sia la densità di X; è invece chiaro che*

$$f_{X|N}(x \mid n) = \frac{1}{\sqrt{2\pi}} e^{-(x-n)^2/2}.$$

Diciamo che $f_{X|N}(x \mid n)$ è la densità di X condizionata ad N.

FRAME 6.11. *Il nostro esperimento si suddivide in due fasi indipendenti: nella prima si decide l'ora n di partenza, nella seconda si misura il tempo necessario. Applicando il principio di moltiplicazione possiamo definire la densità congiunta di N ed X:*

$$h_{N,X}(n,x) = \lim_{\Delta \to 0} \frac{1}{\Delta} P\{x \le X \le x + \Delta, N = n\} =$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta} P\{x \le X \le x + \Delta \mid N = n\} P\{N = n\} =$$

$$= P\{N = n\} \lim_{\Delta \to 0} \frac{1}{\Delta} P\{x \le X \le x + \Delta \mid N = n\} = p_N(n) f_{X|N}(x \mid n).$$

FRAME 6.12. *Dalla densità congiunta ricaviamo facilmente le densità marginali:*

$$p_N(n) = \int_{-\infty}^{\infty} h_{N,X}(n,x)dx = \int_{-\infty}^{\infty} p_N(n) f_{X|N}(x \mid n)dx =$$

$$= p_N(n) \int_{-\infty}^{\infty} f_{X|N}(x \mid n)dx,$$

e, di maggior interesse,

$$f_X(x) = \sum_n h_{N,X}(n,x) = \sum_n p_N(n) f_{X|N}(x \mid n).$$

6.2. **Valore atteso di una funzione di due variabili aleatorie.**

FRAME 6.13. *Siano X ed Y variabili aleatorie discrete aventi densità congiunta p; data una funzione $g : \mathbb{R}^2 \to \mathbb{R}$, si ha*

$$E[g(X,Y)] = \sum_y \sum_x g(x,y)p(x,y).$$

Analogamente, se X ed Y sono variabili aleatorie continue aventi densità congiunta f, data una funzione $g : \mathbb{R}^2 \to \mathbb{R}$, si ha

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy.$$

FRAME 6.14. *Possiamo ora dimostrare che $E[X+Y] = E[X] + E[Y]$:*

$$E[X+Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y)f(x,y)dxdy =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x,y)dydx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x,y)dxdy =$$

$$= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y)dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x,y)dx \right] dy =$$

$$\int_{-\infty}^{\infty} xf_X(x)dx + \int_{-\infty}^{\infty} yf_Y(y)dy.$$

FRAME 6.15. *Siano X_1, ..., X_n variabili aleatorie indipendenti identicamente distribuite. La variabile aleatoria*

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

è detta la media campionaria *di X_1, ..., X_n.*

FRAME 6.16. *Il valore atteso di $\overline{X}$ è*

$$E[\overline{X}] = E\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n} nE[X_1] = E[X_1].$$

La varianza di $\overline{X}$ è

$$Var(\overline{X}) = Var\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n} Var(X_1).$$

6.3. **Covarianza, varianza e correlazioni.**

FRAME 6.17. ***Proposizione*** *Se X ed Y sono variabili aleatorie indipendenti, ed h e g sono due funzioni di una variabile, allora*

$$E[g(X)h(Y)] = E[g(X)]E[g(y)].$$

Dimostrazione. *Supponiamo X ed Y continue con densità congiunta f. Denotiamo con $\ell : \mathbb{R}^2 \to \mathbb{R}$ la funzione $\ell(x, y) = g(x)h(y)$. Allora*

$$E[g(X)h(Y)] = E[\ell(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ell(x, y)f(x, y)dxdy$$

FRAME 6.18.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ell(x, y)f(x, y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x, y)dxdy =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy =$$

$$= \int_{-\infty}^{\infty} h(y)f_Y(y)dy \int_{-\infty}^{\infty} g(x)f_X(x)dx = E[h(X)]E[g(Y)]$$

FRAME 6.19. *In generale date due variabili aleatorie X ed Y, il valore atteso $E[XY]$ non è uguale a $E[X]E[Y]$; diciamo* covarianza *il numero*

$$Cov(X, Y) = E[XY] - E[X]E[Y].$$

Ovviamente la covarianza di due variabili indipendenti è zero.

FRAME 6.20. *Non vale il viceversa! Infatti se X è la variabile aleatoria discreta che assume i soli valori -1, 0, 1 con densità $p(-1) = p(0) = p(1) = \frac{1}{3}$ e Y è la variabile aleatoria*

$$Y = \begin{cases} 0 & se\ X \neq 0, \\ 1 & se\ X = 0, \end{cases}$$

si ha $E[XY] = E[0] = 0 = E[X] = E[X]E[Y]$.

FRAME 6.21. *La covarianza di X ed Y puó essere anche interpretata come una "varianza in due variabili": posto $\mu = E[X]$ e $\nu = E[Y]$ si ha*

$$E[(X - \mu)(Y - \nu)] = E\left[(XY - \mu Y - X\nu + \mu\nu)\right] =$$

$$= E[XY] - \mu E[Y] - E[X]\nu + \mu\nu = E[XY] - E[X]E[Y] = Cov(X, Y).$$

FRAME 6.22.
- *$Cov(X, Y) = Cov(Y, X)$;*
- *$Cov(X, X) = Var(X)$;*
- *$Cov(aX, Y) = aCov(X, Y)$;*

- $Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_i\right) = \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j)$.

Verifichiamo l'ultima, le altre essendo facili conseguenze di risultati precedenti. Sia $\mu_i = E[X_i]$ *e* $\nu_j = E[Y_j]$. *Allora*

$$E[\sum_{i=1}^n X_i] = \sum_{i=1}^n \mu_i, \ \ e\ E[\sum_{j=1}^m Y_j] = \sum_{j=1}^m \nu_j.$$

FRAME 6.23. *Allora*

$$Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_i\right) = E\left[\left(\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i\right)\left(\sum_{j=1}^m Y_j - \sum_{j=1}^m \nu_j\right)\right] =$$

$$= E\left[\sum_{i=1}^n (X_i - \mu_i) \sum_{j=1}^m (Y_j - \nu_j)\right] =$$

$$= E\left[\sum_{i=1}^n \sum_{j=1}^m (X_i - \mu_i)(Y_j - \nu_j)\right] =$$

$$= \sum_{i=1}^n \sum_{j=1}^m E[(X_i - \mu_i)(Y_j - \nu_j)].$$

FRAME 6.24. *Avevamo visto che se due variabili* X, Y *sono indipendenti, allora*

$$Var(X+Y) = Var(X) + Var(Y).$$

Se non sono indipendenti ...

$$Var(X+Y) = Cov(X+Y, X+Y) =$$
$$= Cov(X,X) + 2Cov(X,Y) + Cov(Y,Y) =$$
$$= Var(X) + Var(Y) + 2Cov(X,Y).$$

In generale

$$Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j).$$

6.4. **Valore atteso condizionato.**

FRAME 6.25. *Siano* X *ed* Y *variabili aleatorie discrete. Abbiamo definito la densità condizionata*

$$p_{X|Y}(x \mid y) = P\{X = x \mid Y = y\} = \frac{p(x,y)}{p_Y(y)}.$$

Qual'è la media di X, *dato* $Y = y$*?*

$$E[X \mid Y = y] = \sum_x x p_{X|Y}(x \mid y).$$

Analogamente se X ed Y sono variabili aleatorie continue e $f_Y(y) \neq 0$,

$$E[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y) dx =$$
$$= \int_{-\infty}^{\infty} x \frac{f(x,y)}{f_Y(y)} dx.$$

FRAME 6.26. ***Esempio*** *Siano X ed Y variabili aleatorie continue con densità congiunta*

$$f(x,y) = \frac{e^{-x/y} e^{-y}}{y} \qquad 0 < x < \infty, 0 < y < \infty.$$

Calcolare $E[X \mid Y = y]$.

FRAME 6.27. *Date due variabili aleatorie X ed Y, consideriamo la funzione*

$$g : \mathbb{R} \to \mathbb{R}, \quad y \mapsto E[X \mid Y = y].$$

La composizione $g \circ Y$ è una variabile aleatoria che indichiamo con $E[X \mid Y]$. Se Y è discreta (continua) anche $E[X \mid Y]$ è discreta (continua). Quanto vale il suo valore atteso?

FRAME 6.28. ***Proposizione*** $E[E[X \mid Y]] = E[X]$.

Se Y è discreta, si ha pertanto

$$E[X] = E[g(Y)] = \sum_y E[X \mid Y = y] P\{Y = y\};$$

Se Y è continua, si ha pertanto

$$E[X] = E[g(Y)] = \int_{-\infty}^{\infty} E[X \mid Y = y] f_Y(y) dy.$$

FRAME 6.29. *Un minatore è bloccato in una galleria dove ci sono tre porte. La prima porta conduce ad un tunnel che in 3 ore porta il minatore in salvo; la seconda porta conduce ad un tunnel che in 5 ore lo fa accedere ad una nuova galleria con 3 porte analoghe a quelle della prima galleria; la terza porta conduce ad un tunnel che in 7 ore lo fa accedere ad una nuova galleria con 3 porte analoghe a quelle della prima galleria. Determinare il valore atteso del tempo impiegato dal minatore per salvarsi se ogni volta sceglie a caso una delle porte.*

FRAME 6.30. *Sia A un evento; definiamo la variabile aleatoria*

$$X = \begin{cases} 1 & \text{se } A \text{ si realizza}, \\ 0 & \text{altrimenti}. \end{cases}$$

Chiaramente X è una variabile di Bernoulli di parametro $P(A)$ e quindi di valore atteso $E[X] = P(A)$.

Sia Y un'altra variabile aleatoria. Allora

$$E[X \mid Y = y] = P(A \mid Y = y)$$

è il valore atteso di X quando restringiamo lo spazio campionario a $Y = y$, ovvero a $Y^{-1}(y)$.

FRAME 6.31. *Pertanto se Y è discreta si ha*

$$P(A) = E[X] = E[E[X \mid Y]] = \sum_y E[X \mid Y = y]P\{Y = y\} =$$

$$= \sum_y P(A \mid Y = y)P\{Y = y\};$$

se invece Y è continua

$$P(A) = E[X] = E[E[X \mid Y]] = \int_{-\infty}^{\infty} E[X \mid Y = y]f(y)dy =$$

$$= \int_{-\infty}^{\infty} P(A \mid Y = y)f(y)dy.$$

FRAME 6.32. *Siano X ed Y variabili aleatorie continue indipendenti con densità f_X e f_Y. Calcolare $P\{X < Y\}$. Determinare la distribuzione di $X + Y$.*