3.33pt

Psicometria 1 (023-PS)

Michele Grassi
mgrassi@units.it

Università di Trieste

Lezione 8 9

Piano della presentazione

3.33pt

Variabili aleatorie discrete

Variabili aleatorie discrete

- Una variabile aleatoria è una regola che assegna un numero a ciascun evento definito sullo spazio campione di un esperimento aleatorio.
- La variabile aleatoria X, per esempio, corrisponde al numero di eventi T nell'esperimento che consiste nel lancio di due monete.

Esito	Valore x di X
TT	2
TC	1
CT	1
CC	0

Variabili aleatorie discrete

- La variabile aleatoria Y, per esempio, corrisponde al numero di eventi T al primo lancio nell'esperimento (aleatorio) che consiste nel lancio di due monete.

Esito	Valore y di Y
TT	1
TC	1
CT	0
CC	0

Variabili aleatorie discrete

- Talvolta è utile distinguere tra la variabile aleatoria X (denotata da una lettera maiuscola) e il particolare valore x che essa assume (denotata da una lettera minuscola).

Distribuzione di probabilità

- La distribuzione di probabilità di una variabile aleatoria discreta elenca tutti i possibili valori x_i della variabile, insieme alla probabilità p_i di osservare ciascun valore.
- (Es. 1) La distribuzione di probabilità della variabile X che corrisponde al numero di eventi T nel lancio di due monete oneste, per esempio, è

Esito	x_i	$p_i = P(X = x_i)$
CC	0	0.25
TC, CT	1	0.50
TT	2	0.25
	somma	1

Distribuzione di probabilità

- Si noti che gli eventi semplici dello spazio campione sono equiprobabili, ma le probabilità associate ai valori della variabile aleatoria X non sono uniformi.
- (Es. 2) La distribuzione di probabilità della variabile Y che corrisponde al numero di eventi T *nel primo lancio* di due monete oneste, per esempio, è

Esito	y_i	$p_i = P(Y = y_i)$
CC, CT	0	0.50
TC, TT	1	0.50
	somma	1

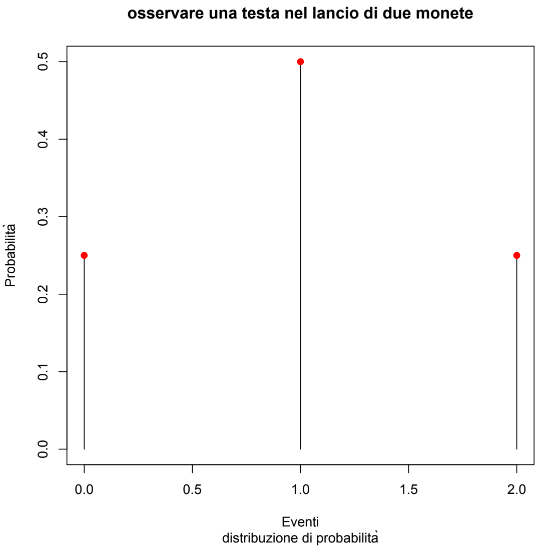
Distribuzione di probabilità

- In generale, una variabile aleatoria discreata ha un numero finito di valori possibili, $x_1, x_2, \ldots, x_k$, con probabilità $p_1, p_2, \ldots, p_k$.
  - In base agli assimi della teoria della probabilità, ciascun p_i è un valore compreso tra 0 e 1, e la somma di tutte le probabilità è $p_1 + p_2 + \ldots + p_k = 1.0$.
  - La probabilità di un particolare evento riferito alla variabile aleatoria si calcola sommando le probabilità associate a tutti i valori che costituiscono l'evento. La probabilità di ottenere almeno una volta testa, per esempio, è $P(X = 1) + P(X = 2) = 0.50 + 0.25 = 0.75$.

La probabilità di una variabile aleatoria discreta può essere rappresentata con un grafico.

```
esiti<-c(0,1,2)
prob<-c(0.25,0.5,0.25)
plot(esiti,prob,type="h",ylim=c(0,.5),
      sub="distribuzione di probabilità",
      main="osservare una testa nel lancio di due monete",
      xlab="Eventi",
      ylab="Probabilità")
points(x=esiti,y=prob,pch=19,col="red")
```

Distribuzione di probabilità



Media, Varianza e Deviazione Standard

Media

- La media μ di una variabile aleatoria X, detta valore atteso, $E(X)$, è definita nel modo seguente:

$$\mu = x_1p_1 + x_2p_2 + \ldots + x_kp_k = \sum_{i=1}^{k} x_ip_i$$

- Per l'esempio precedente avremo:

	x_i	p_i	x_ip_i
CC	0	0.25	0.00
TC, CT	1	0.50	0.50
TT	2	0.25	0.50
	somma	1	$\mu = 1$

Interpretazione del valore atteso

Il valore atteso di X è la media aritmetica dei valori della variabile aleatoria nel senso seguente.

- Il valore atteso $E(X)$ è la media di tutti i valori che X può assumere, ciascuno pesato dalla sua probabilità.
- Se pensiamo alle probabilità come a dei pesi attaccati ad una barra, allora il valore atteso $E(X)$ rappresenta il punto di equilibrio (centro di massa)
- Se l'esperimento viene ripetuto molte volte e il valore di X viene calcolato in ciascuna prova, allora la media di questi valori di X sarà tanto più simile a $E(X)$ quante più volte l'esperimento verrà ripetuto.

Interpretazione del valore atteso

Il valore atteso $E(X)$ è la media di tutti i valori che X può assumere, ciascuno pesato dalla sua probabilità.

- Consideriamo la formula della statistica $\bar{x}$ −*media (aritmetica) campionaria*:

$$\bar{x} = \sum_{i=1}^{n} x_i/n,$$

che è la somma di n osservazioni della variabile X, ciascuna pesata per $1/n$.

- Non conoscendo il modello probabilistico per una variabile osservata x, la migliore stima della probabilità di comparsa è $1/n$, ossia una distribuzione uniforme.
- Se nel campione compariranno due valori uguali (es., $x_1 = x_2 = 2.56$), allora nella formula della media otterremo che il relativo peso diventerà $2/n$ (infatti $x_1/n + x_2/n = 2.56/n + 2.56/n = 2.56 \cdot 2/n$).

Interpretazione del valore atteso

Il valore atteso $E(X)$ rappresenta il centro di massa (baricentro)

- Essendo $\sum p_i = 1$, possiamo quindi riscrivere $E(X)$ come

$$E(X) = \mu = \frac{\sum_{i=1}^{k} x_i p_i}{\sum p_i}$$

  che rappresenta la *coordinata del centro di massa di un sistema di punti*, ciascuno avente una "massa di probabilità" p_i.
- Con una espressione mutuata dalla *meccanica*, possiamo affermare che $E(X)$ rappresenta il baricentro della distribuzione di probabilità. Questa è un' ulteriore giustificazione dell'uso di $E(X)$ per quantificare il valore intorno al quale ci aspettiamo che la variabile casuale assuma il valore che verrà osservato.
- Il valore atteso non corrisponde, in generale, né al valore più probabile, né a uno dei possibili valori che la variabile casuale può assumere.

Interpretazione del valore atteso

Se l'esperimento viene ripetuto molte volte e il valore di X viene calcolato in ciascuna prova, allora la media di questi valori di X sarà tanto più simile a $E(X)$ quante più volte l'esperimento verrà ripetuto.

- come abbiamo detto a riguardo dell'interpretazione frequentista della probabilità, le frequenze relative $f_i = n_i/n$, si approssimano sempre di più alle probabilità p_i al crescere del numero di ripetizioni dell'esperimento.
- Ripetere molte volte l'esperimento e calcolare la frequenza relativa con cui l'evento di interesse si è verificato è una procedura empirica per stimare le probabilità.
- Con n tendente all'infinito, $n \to \infty$,

$$\bar{x} = \sum_{i=1}^{n} x_i \frac{1}{n} \to \sum_{i=1}^{k} x_i \frac{n_i}{n} \to = \mu = \sum_{i=1}^{k} x_i p_i$$

Varianza

Definizione

Sia X una variabile aleatoria con funzione di probabilità $p(x)$ e valore atteso μ. Allora, la varianza di X, denotata da $V(X)$ o σ_X^2 o soltanto σ^2, è

$$V(X) = \sum_{i=1}^{k} (x_i - \mu)^2 \cdot p(x_i) = E\big[(X - \mu)^2\big]$$

Formula alternativa

$$V(X) = \sigma^2 = \sum_{i=1}^{k} \left[x_i^2 \cdot p(x_i)\right] - \mu^2 = E\left(X^2\right) - \left[E(X)\right]^2$$

sviluppiamo la formula originale

$$\begin{aligned}
V(X) &= \sum (x_i - \mu)^2 \cdot p(x_i) \\
&= \sum (x_i^2 + \mu^2 - 2x_i\mu) \cdot p(x_i) \\
&= \sum \left(x_i^2 \cdot p(x_i)\right) + \sum \left(\mu^2 \cdot p(x_i)\right) - 2\sum \left(x_i\mu \cdot p(x_i)\right) \\
&= E\left(X^2\right) + \mu^2 \cdot \sum p(x_i) - 2\mu \sum \left(x_i \cdot p(x_i)\right) \\
&= E\left(X^2\right) + \mu^2 \cdot 1 - 2\mu^2 \\
&= E\left(X^2\right) - \mu^2 = E\left(X^2\right) - \left[E(X)\right]^2
\end{aligned}$$

Varianza

Proprietà

$$V(aX+b) = a^2 \cdot \sigma_X^2$$

da cui l'implicazione che $\sigma_{X+b}^2 = \sigma_X^2$

sviluppiamo con la formula alternativa di $V(X)$

$$V(aX+b) = E\left[(aX+b)^2\right] - \left[E(aX+b)\right]^2$$

sapendo che $E(aX+b) = \sum_{i=1}^{k}(ax_i+b)p(x_i)$,

$$\begin{aligned}
V(aX+b) &= E\left[(a^2X^2+b^2+2abX)\right] - \left[aE(X)+b\cdot 1)\right]^2 \\
&= a^2E\left[X^2\right] + b^2 + 2abE(X) - a^2\left[E(X)\right]^2 - b^2 - 2abE(X) \\
&= a^2\left\{E\left[X^2\right] - \left[E(X)\right]^2\right\} + b^2 - b^2 + 2abE(X) - 2abE(X) \\
&= a^2\sigma_X^2
\end{aligned}$$

Definizione

La deviazione standard di X è

$$\sigma_X = \sqrt{\sigma_X^2}$$

- Per l'esempio precedente avremo:

	x_i	p_i	$x_i p_i$	$x_i - \mu$	$(x_i - \mu)^2 p_i$
CC	0	0.25	0.00	-1	0.25
TC, CT	1	0.50	0.50	0	0.00
TT	2	0.25	0.50	1	0.25
	somma:	1	$\mu = 1$		$\sigma^2 = 0.50$

$$\sigma_X = \sqrt{0.50} \approx 0.71$$

Esempio

- Consideriamo la variabile aleatoria discreta X avente la seguente distribuzione di probabilità

x_i	p_i
1	0.2
2	0.8
somma	1

Illustrazione

- Possiamo pensare alla variabile X può nel modo seguente:
  - l'esperimento consiste nell'estrazione casuale con rimessa (*infinite volte*) di una pallina da un'urna contenente 10 palline numerate nel modo seguente: $\Omega = \{1, 1, 2, 2, 2, 2, 2, 2, 2, 2\}$;
  - l'esperimento consiste nell'estrazione casuale senza rimessa di una pallina da un'urna contenente *infinite palline*; il 20% di tali palline porta il numero 1 il rimanente 80% il numero 2.

Illustrazione

- Nei termini della prima delle due rappresentazioni (equivalenti) descritte in precedenza è immediatamente evidente che media e varianza di X si calcoleranno sui 10 numeri $\Omega = \{1; 1; 2; 2; 2; 2; 2; 2; 2; 2\}$ attraverso le formule utilizzate dalla statistica descrittiva.
- I pesi $1/n$ associati a ciascuna osservazione coincideranno con le $p(x_i)$ teoriche note di realizzo nell'esperimento, e quindi le formule di media e varianza coincideranno con i calcoli di valore atteso e varianza di una variabile aleatoria discreta.

Illustrazione

- Il valore atteso di X è

$$E(X) = \sum_{i=1}^{2} x_i p_i = x_1 p_1 + x_2 p_2 = 1 \times 0.2 + 2 \times 0.8 = 1.8$$

- La varianza di X è

$$\begin{aligned} V(X) &= \sum_{i=1}^{2} \left(x_i - \mu\right)^2 p_i \\ &= (1 - 1.8)^2 \times 0.2 + (2 - 1,8)^2 \times 0.8 = 0.16 \\ &= (0.4)^2 \end{aligned}$$

Illustrazione

- Gli stessi risultati si ottengono con le formule per il calcolo di media e varianza della statistica descrittiva, trattando lo spazio campionario Ω coma una *variabile osservata su un campione di ampiezza* $n = 10$.

```
 x<-c(1,1,2,2,2,2,2,2,2,2)
 x
## [1] 1 1 2 2 2 2 2 2 2 2
 mean(x)
## [1] 1.8
 var(x)*9/10
## [1] 0.16
 sd(x)*sqrt(9)/sqrt(10)
## [1] 0.4
```

- Si noti il denominatore della varianza $n = 10$, rispetto ai $n - 1 = 9$ gradi di libertà. Tale correzione, come detto, è necessaria nella formula della statistica campionaria s^2 e non nel parametro σ^2.

Illustrazione

- sviluppando la media campionaria degli (n=10) valori X abbiamo:

$$
\begin{aligned}
\bar{\mathbf{x}} = &\quad 1 \times \frac{1}{10} + 1 \times \frac{1}{10} + 2 \times \frac{1}{10} + 2 \times \frac{1}{10} + 2 \times \frac{1}{10} + 2 \times \frac{1}{10} + \\
&\quad + 2 \times \frac{1}{10} + 2 \times \frac{1}{10} + 2 \times \frac{1}{10} + 2 \times \frac{1}{10} \\
= &\quad 1 \times \left(\frac{1}{10} \times 2\right) + 2 \times \left(\frac{1}{10} \times 8\right) \\
= &\quad 1 \times 0.2 + 2 \times 0.8 = \mathbf{E}(\mathbf{X}) = \mathbf{1.8}
\end{aligned}
$$

- sviluppando la varianza campionaria abbiamo:

$$
\begin{aligned}
\mathbf{s^2_x} = &\quad (1-1.8)^2\frac{1}{10} + (1-1.8)^2\frac{1}{10} + (2-1.8)^2\frac{1}{10} + (2-1.8)^2\frac{1}{10} + \\
&\quad + (2-1.8)^2\frac{1}{10} + (2-1.8)^2\frac{1}{10} + (2-1.8)^2\frac{1}{10} + \\
&\quad + (2-1.8)^2\frac{1}{10} + (2-1.8)^2\frac{1}{10} + (2-1.8)^2\frac{1}{10} \\
= &\quad (1-1.8)^2 \times \left(\frac{1}{10} \times 2\right) + (2-1.8)^2 \times \left(\frac{1}{10} \times 8\right) \\
= &\quad (1-1.8)^2 0.2 + (2-1.8)^2 0.8 = \mathbf{E}\left[(\mathbf{X}-\mu)^{\mathbf{2}}\right] = \mathbf{0.16}
\end{aligned}
$$

Interpretazione

- Che cos'è allora che rende X una variabile aleatoria?
  - L'esperimento considerato consiste nell'estrazione casuale con rimessa di una pallina dall'urna $\Omega = \{1; 1; 2; 2; 2; 2; 2; 2; 2; 2\}$. Dunque, il valore x_i ottenuto in ciascuna singola estrazione non può essere previsto a priori:
    - il valore della variabile aleatoria X non può essere conosciuto prima di avere eseguito l'esperimento;

Interpretazione

- Possiamo però sapere cosa succede se un grande numero di estrazioni vengono effettuate (se un grande numero di valori della variabile aleatoria viene osservato):
  - media e varianza di un grande numero di estrazioni diventeranno sempre più simili a μ e σ_x^2 al crescere del numero di ripetizioni dell'esperimento (al crescere del numero di valori di X osservati).

Interpretazione

- Consideriamo un caso concreto di questo processo.
- Utilizzando R, generiamo una sequenza di valori della variabile (aleatoria) X, avente la seguente distribuzione di probabilità:

x_i	p_i
1	0.2
2	0.8
somma	1

Interpretazione

$N = 10$

```
 s_1<-sample(x=c(1,2),size=10,replace=TRUE,prob=c(0.2,0.8))
 mean(s_1)
# [1] 1.9
 s_1<-sample(x=c(1,2),size=10,replace=TRUE,prob=c(0.2,0.8))
 mean(s_1)
# [1] 1.9
```

$N = 50$

```
 s_1<-sample(x=c(1,2),size=50,replace=TRUE,prob=c(0.2,0.8))
 mean(s_1)
# [1] 1.8
 s_1<-sample(x=c(1,2),size=50,replace=TRUE,prob=c(0.2,0.8))
 mean(s_1)
# [1] 1.74
```

Interpretazione

$N = 10$

```
 s_1<-sample(x=c(1,2),size=10,replace=TRUE,prob=c(0.2,0.8))
 var(s_1)*9/10
# [1] 0.09
 s_1<-sample(x=c(1,2),size=10,replace=TRUE,prob=c(0.2,0.8))
 var(s_1)*9/10
# [1] 0.21
```

$N = 50$

```
 s_1<-sample(x=c(1,2),size=50,replace=TRUE,prob=c(0.2,0.8))
 var(s_1)*49/50
# [1] 0.1824
 s_1<-sample(x=c(1,2),size=50,replace=TRUE,prob=c(0.2,0.8))
 var(s_1)*49/50
# [1] 0.1716
```

Le formule per μ, σ^2 e σ sono molto simili alle formule per la media, varianza e deviazione standard di un campione di osservazioni.

	variabile aleatoria	variabile in un campione
media	$\mu = \sum x_i \cdot p_i(x_i)$	$\bar{x} = \frac{1}{n} \sum x_i$
varianza	$\sigma^2 = \sum (x_i - \mu)^2 \cdot p(x_i)$	$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$
deviazione standard	$\sqrt{\sigma^2}$ ciao	$\sqrt{s^2}$

nota: il denominatore di $s^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$, n, diventa $n - 1$ ***gradi di libertà*** per ragioni che vedremo più in là.

Funzione cumulativa di probabilità

Definizione

La funzione cumulativa di probabilità (*cdf*) $F(x)$ di una variabile aleatoria discreta X con funzione di probabilità $p(x)$ è

$$F(x) = P(X \leq x) = \sum_{y|y \leq x} p(y)$$

esempio

Le *cdf* vengon usate per calcolare le probabilità. Per l'esempio relativo al lancio di due monete, avremo

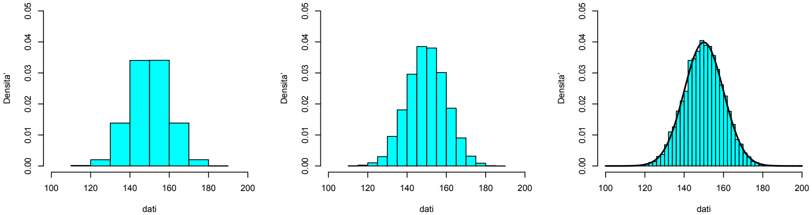
$$P(X \leq 1) = P(X = 0) + P(X = 1) = 0.25 + 0.50 = 0.75$$

Variabili aleatorie continue

Variabili aleatorie continue

- Le variabili aleatorie definite su uno spazio campione continuo possono essere a loro volta continue
- La distribuzione di probabilità di una variabile aleatoria continua è descritta da una curva (funzione) di densità, $p(x)$.
  - Non ha senso parlare della probabilità di osservare uno specifico, particolare valore di una variabile aleatoria continua.
  - Invece, l'area sottese alla curva di densità in un dato intervallo di valori corrisponde alla probabilità di osservare un valore della variabile aleatoria compreso in quell'intervallo.

Significato La funzione di densità può essere considerata come il limite a cui tende un istogramma discreto.



Esempio

Sia X la misura dell'altezza delle persone.

- Possiamo rendere discreta X misurando l'altezza in centimetri, oppure in millimetri.
- Tanto più precise saranno le nostre misurazioni, tanto più l'istogramma che potremmo produrre avendo a disposizione un grande numero di misure si approssimerà ad una curva continua.
- Questa curva continua è la funzione di densità di probabilità (fdp).

Variabili aleatorie continue

- Le variabili aleatorie continue, allo stesso modo delle variabili discrete, hanno una media, varianza e deviazione standard.
- Le formule per la media e la varianza di una variabile aleatoria continua sono simili alle formule corrispondenti per le variabili discrete (sostituendo gli integrali alle somme).

$$\mu = E(X) = \int_{-\infty}^{\infty} x p(x)\,\mathrm{d}x,$$

$$\sigma_x^2 = V(X) = E\big[(X-\mu)^2)\big] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\,\mathrm{d}x.$$

E la deviazione standard di X

$$\sigma_x = \sqrt{V(X)}$$

Funzione cumulativa di densità

- La funzione cumulativa di densità $F(x)$ di una variabile aleatoria continua X per ogni numero x è definita da

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(x)\,\mathrm{d}x$$

- Per ciascun x, $F(x)$ è l'area sottesa alla curva alla sinistra di x.

Calcolo delle probabilità con la fcd.

- Sia X una variabile aleatoria continua con fdp $p(x)$ e fcp $F(x)$. Allora, per ogni numero a,

$$P(X > a) = 1 - F(a)$$

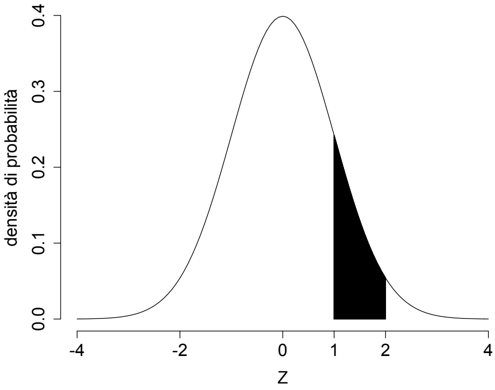
e per ogni coppia di numeri a e b, con $a < b$,

$$P(a \leq X \leq b) = F(b) - F(a)$$

- Nella figura seguente è rappresentata l'area sottesa alla funzione densità di probabilità *normale standardizzata* nell'intervallo compreso tra 1 e 2. Tale area corrisponde alla probabilità $P(1 \leq Z \leq 2)$.

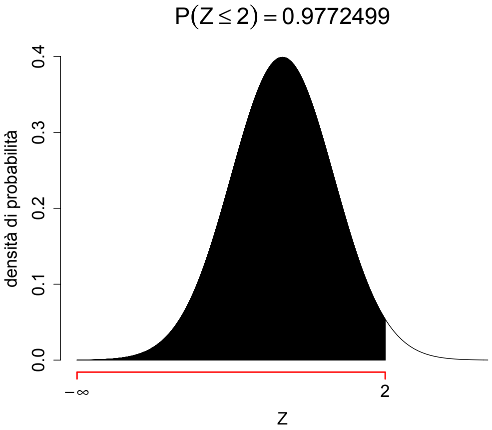
Distribuzione normale standardizzata

$P(1 \leq Z \leq 2) = F(2.0) - F(1.0) =$
$0.9772 - 0.8413 = 0.1359$ per $Z \sim N(0; 1)$.



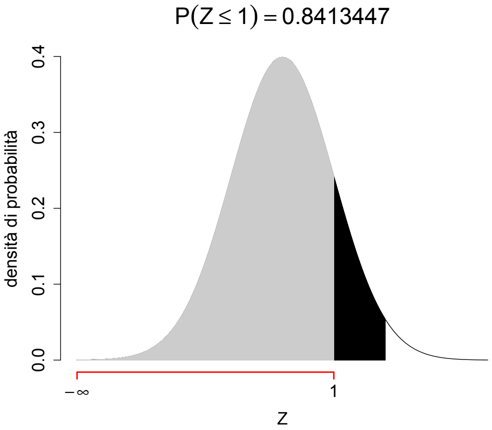
Distribuzione normale standardizzata

$P(1 \leq Z \leq 2) = F(2.0) - F(1.0) =$
$\boxed{0.9772}$.

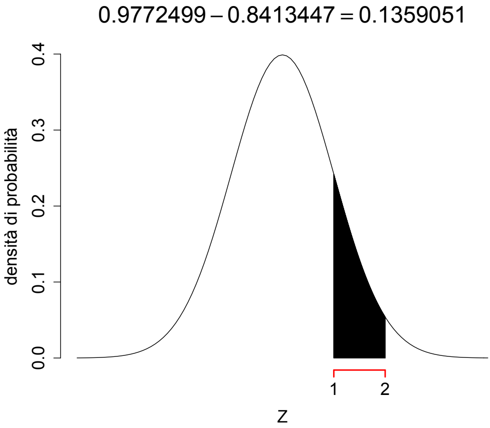


Distribuzione normale standardizzata

$P(1 \leq Z \leq 2) = F(2.0) - F(1.0) =$
$0.9772 - \boxed{0.8413}$.



$P(1 \leq Z \leq 2) = F(2.0) - F(1.0) =$
0.9772-0.8413=0.1359 per $Z \sim N(0; 1)$.



Distribuzioni binomiali

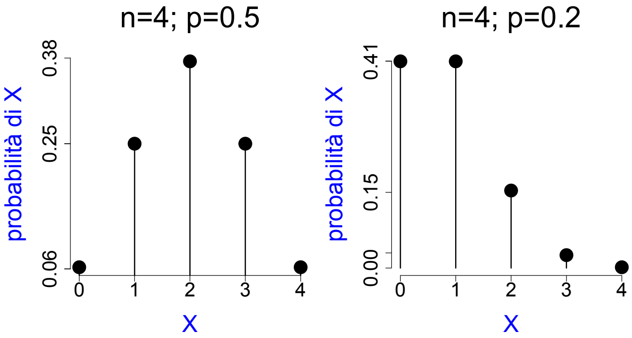
Distribuzioni binomiali

- Le distribuzioni binomiali sono una famiglia di distribuzioni di probabilità.
- Una variabile aleatoria binomiale X si osserva nelle seguenti condizioni:
  - c'è un numero fisso di n osservazioni;
  - le osservazioni sono *indipendenti* le une dalle altre;
  - ciascuna osservazione corrisponde ad uno di due esiti possibili, convenzionalmente chiamati "**successo**" e "**insuccesso**";
  - la probabilità di un successo, denotata da p, rimane costante per tutte le osservazioni.

Distribuzioni binomiali

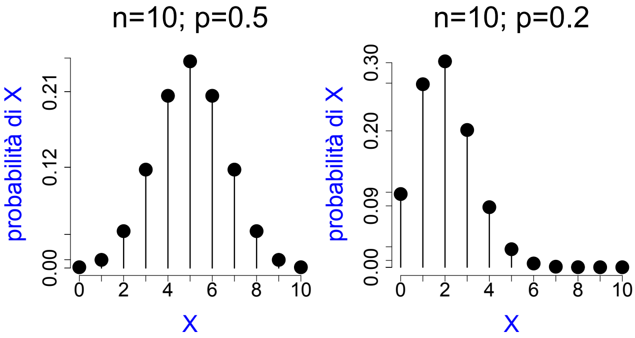
- La variabile X corrisponde al numero di successi in n prove e dunque è un numero compreso tra 0 e n.
- La distribuzione binomiale con parametri n e p fornisce l'elenco delle probabilità associate a ciascun possibile valore X.
- Alcune distribuzioni binomiali sono rappresentate graficamente nelle figure seguenti.

Distribuzione binomiali



Numero di successi in $n = 4$ esperimenti

Distribuzione binomiali



Numero di successi in $n = 10$ esperimenti

Distribuzioni binomiali

- Consideriamo l'esperimento che consiste nel lanciare due monete oneste e nel contare il numero X di eventi "testa".
- Lo spazio campione dell'esperimento è $\Omega = \{TT, TC, CT, CC\}$ e ciascun evento nello spazio campione ha probabilità $1/4$.
- La distribuzione di probabilità di X è

x_i	p_i
0	$1/4$
1	$1/2$
2	$1/4$
somma	1

- **Questa è la distribuzione binomiale per** $n = 2$ **e** $p = 0.5$.

Distribuzioni binomiali

- **Consideriamo ora la distribuzione binomiale con** $n = 4$ **e** $p = 0.2$. [cfr. Figura]
- Possiamo immaginare questa distribuzione come quella che descrive l'esperimento in cui vengono lanciate 4 monete *disoneste*, tali per cui la probabilità di osservare "testa" è 0.2.
- Come in precedenza, contiamo il numero di eventi "testa" effettivamente osservati.

- Per calcolare la distribuzione di X potremmo elencare tutti gli eventi dello spazio campione Ω, contando gli eventi (sul totale) che soddisfano ciascuna modalità di X.
    - In questo caso lo spazio campione contiene $2 \times 2 \times 2 \times 2 = 2^4 = 16$ eventi.
    - Questa strategia (*punto campione*), però, diventa difficilmente praticabile al crescere di n.
    - Possiamo invece usare un metodo diverso. Consideriamo, ad esempio, $P(X = 2)$, ovvero la probabilità dell'evento che consiste nell'osservare due esiti "testa" in 4 lanci
-

- Due teste vengono osservate, per esempio, nella sequenza $TTCC$. Dato che i lanci sono indipendenti, tale evento ha una probabilità di

$$P(TTCC) = 0.2 \times 0.2 \times 0.8 \times 0.8 = 0.2^2 \times 0.8^2 = 0.0256$$

.

- Due teste vengono osservate, però, anche nella sequenza $TCTC$, a cui associamo la stessa probabilità:

$$P(TCTC) = 0.2 \times 0.8 \times 0.2 \times 0.8 = 0.2^2 \times 0.8^2 = 0.0256$$

Distribuzioni binomiali

- Qualunque sequenza contenente 2 teste e 2 croci, non importa in quale ordine, avrà probabilità

$$0.2^2 \times 0.8^2 = 0.0256.$$

- In quanti modi diversi si possono ottenere 2 teste in 4 lanci?
- Dato che lo spazio campione è piccolo, possiamo semplicemente contare tutte le combinazioni:

$$TTCC \;\; CTTC \;\; CCTT \; TCCT \; TCTC \;\; CTCT$$

Distribuzioni binomiali

- In conclusione, la probabilità di ottenere due teste in 4 lanci è

$$
\begin{aligned}
P(X=2) = & \; P(TTCC \cup CTTC \cup CCTT \cup TCCT \cup TCTC \cup CTCT) \\
= & \; 0.0256 + 0.0256 + 0.0256 + 0.0256 + 0.0256 + 0.0256 \\
= & \; 6 \times 0.0256 \\
= & \; \text{combinazioni} \times \text{prob. di ogni sequenza} = 0.1536
\end{aligned}
$$

Distribuzioni binomiali

- Esaminiamo ora una regola generale, basata sul calcolo combinatorio, che consente di trovare le probabilità della distribuzione binomiale.
- Il numero di possibili combinazioni di k successi e $n - k$ insuccessi è data dal coefficiente binomiale:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

dove $n! = n \times (n-1) \times \ldots \times 3 \times 2 \times 1$. Per convenzione $0! = 1$; quindi

$$\binom{n}{0} = \binom{n}{n} = 1$$

Distribuzioni binomiali

- La probabilità di ciascuna sequenza di k successi e $n-k$ insuccessi in n prove è:

$$p^k(1-p)^{n-k}$$

  - si ricordi che $a^0 = 1$ per ciascun numero $a \neq 0$.
  - Se k o $n-k$ sono uguali a 0, dunque, il fattore corrispondente p^0 o $(1-p)^0$ è 1.

Distribuzioni binomiali

- Utilizzando i due risultati precedenti, otteniamo la formula per calcolare le probabilità della distribuzione binomiale:

$$
\begin{aligned}
P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}
\end{aligned}
$$

- Per esempio, la probabilità di ottenere $k = 2$ successi in $n = 4$ prove con probabilità di successo $p = 0.2$ sarà:

$$
\begin{aligned}
P(X=2) = \quad & \binom{4}{2} 0.2^2 (1-0.2)^{4-2} \\
= \quad & \frac{4!}{2!(4-2)!} 0.2^2 0.8^2 \\
= \quad & \frac{4 \times 3 \times 2 \times 1}{(2 \times 1)(2 \times 1)} 0.2^2 0.8^2 \\
= \quad & 0.1536
\end{aligned}
$$

Distribuzione binomiale

- La distribuzione binomiale con parametri $n = 4$ e $p = 0.2$ diventa quindi

k	$P(X = k)$
0	0.4096
1	0.4096
2	0.1536
3	0.0256
4	0.0016
somma	1.0000

```
dbinom(x=0:4,size=4,prob=0.2)
## [1] 0.4096 0.4096 0.1536 0.0256 0.0016
```

Media e varianza

- La distribuzione di Bernoulli o *bernoulliana* è un caso speciale della distribuzione binomiale, dove $n = 1$.
- Simbolicamente, $X \sim B(1, p)$ ha lo stesso significato di $X \sim Bern(p)$.
- Similmente, ciascuna distribuzione binomiale, $B(n, p)$, è la distribuzione della somma di n prove di bernoulli tra loro indipendenti, ciascuna con la stessa probabilità p di *successo*.

- La variabile aleatoria bernoulliana X_bern assume quindi due possibili esiti, 0 (nessun successo) e 1 (successo), con

$$\begin{aligned} P(X_{Bern} = 1) = & \quad p \\ P(X_{Bern} = 0) = & \quad 1 - p. \end{aligned}$$

- la cui media e varianza si ricavano facilmente come

$$\begin{aligned} E(X_{Bern}) = & \quad 1 \times p + 0 \times (1 - p) = p \\ Var(X_{Bern}) = & \quad (1 - p)^2 p + (0 - p)^2 (1 - p) \\ = & \quad (1 - p) \left[(1 - p) p + p^2 \right] \\ = & \quad (1 - p) \left[p - p^2 + p^2 \right] = p(1 - p) \end{aligned}$$

- Come detto, una **variabile aleatoria binomiale** $X \sim B(n,p)$, con parametri n e p, è un processo di Bernoulli, ossia una serie di n variabili aleatorie indipendenti X_i di uguale distribuzione di Bernoulli ($Bern(p)$), dette appunto prove di Bernoulli.
- Le modalità di $X \sim B(n,p)$ sono la somma degli esiti di "successo" (o 1 o 0) in ciascuna prova $X_i \sim Bern(p)$:

$$X_B = X_{Bern\,1} + X_{Bern\,2} + \ldots + X_{Bern\,n}.$$

Media e Varianza

- Perciò, la media e varianza della variabile binomiale X sono presto ricavate:

$$
\begin{aligned}
\mathbf{E(X_B)} = &\quad E(X_{Bern\ 1}) + E(X_{Bern\ 2}) + \ldots + E(X_{Bern\ n}) = \mathbf{np} \\
\mathbf{Var(X_B)} = &\quad Var(X_{Bern\ 1}) + Var(X_{Bern\ 2}) + \ldots + Var(X_{Bern\ n}) \\
= &\quad \mathbf{np(1-p)}
\end{aligned}
$$

Illustrazione

- La distribuzione binomiale con parametri $n = 4$ e $p = 0.2$ avrà quindi

$$\mu = 4 \times 0.2 = 0.8$$

$$\sigma^2 = 4 \times 0.2 \times 0.8 = 0.64$$

$$\sigma = \sqrt{0.64} = 0.8$$

È solo una peculiarità di questo particolare esempio che μ e σ siano uguali.

Illustrazione

```
 samp_1<-rbinom(n=100,size=4,prob=0.2)
 samp_1
##   [1] 0 2 0 2 1 2 2 1 2 0 0 0 1 2 2 0 2 0 1 0 1 1 2 0 0 0
##  [27] 0 1 0 0 1 3 0 1 2 0 1 2 0 1 0 1 0 1 1 0 1 1 0 2 0 0
##  [53] 0 0 3 1 0 0 1 1 0 0 0 3 1 3 0 1 3 1 2 0 0 1 2 1 2 3
##  [79] 0 2 1 1 1 1 1 1 1 1 0 0 1 0 0 0 0 0 0 1 1 1
 mean(samp_1)
## [1] 0.86
  4*.2
## [1] 0.8
 samp_2<-rbinom(n=100,size=4,prob=0.2)
 mean(samp_2)
## [1] 0.68

 samp_3<-rbinom(n=1000,size=4,prob=0.2)
 mean(samp_3)
## [1] 0.812
```

Illustrazione

```
samp_4<-rbinom(n=100,size=4,prob=0.2)
samp_4
##  [1] 0 1 1 2 0 0 1 0 0 1 0 2 0 2 1 1 1 0 1 1 2 0 1 1 0 1
## [27] 0 0 1 1 2 1 1 1 1 1 0 1 0 2 1 1 0 1 0 2 0 0 0 0 0 0
## [53] 0 0 1 0 1 0 1 1 1 2 0 2 2 0 1 2 0 1 0 1 0 0 0 2 0 1
## [79] 2 0 1 0 0 1 0 0 2 0 1 2 1 0 0 1 0 1 1 1 0 1
sqrt(4*0.2*0.8)
## [1] 0.8
sqrt(var(samp_4)*99/100)
## [1] 0.7112665

samp_5<-rbinom(n=100,size=4,prob=0.2)
sqrt(var(samp_5)*99/100)
## [1] 0.7141428

samp_6<-rbinom(n=1000,size=4,prob=0.2)
sqrt(var(samp_6)*999/1000)
## [1] 0.8170135
```

Illustrazione

```
 samp_7<-rbinom(n=10000,size=4,prob=0.2)
 sqrt(var(samp_7)*9999/10000)
## [1] 0.8062912
 mean(samp_7)
## [1] 0.8038

 samp_8<-rbinom(n=10000,size=4,prob=0.2)
 sqrt(var(samp_8)*9999/10000)
## [1] 0.7925528
 mean(samp_8)
## [1] 0.8001

 samp_9<-rbinom(n=10000,size=4,prob=0.2)
 sqrt(var(samp_9)*9999/10000)
## [1] 0.7943017
 mean(samp_9)
## [1] 0.7978
```

Conclusioni

- La distribuzione di probabilità di una variabile aleatoria discreta elenca tutti i possibili valori che la variabile può assumere insieme alla probabilità di ciascun valore.
- La distribuzione di probabilità di una variabile aleatoria continua è rappresentata da una funzione di densità.
- La media di una distribuzione di probabilità è detta valore atteso. - La distribuzione binomiale è la distribuzione di probabilità della variabile aleatoria discreta che rappresenta il numero di successi in n prove bernoulliane.
- Il valore atteso e la varianza di una distribuzione binomiale con parametri n e p sono $\mu = np$ e $\sigma^2 = np(1-p)$.
- le note statistiche campionarie $\bar{x}$ e s^2 tendono a questi valori, al crescere del numero di osservazioni estratte.