

Distribuzioni di probabilità e distribuzione binomiale

Esercizi svolti — variabile uniforme discreta, permutazioni e combinazioni, distribuzione binomiale, test d'ipotesi e test del segno.

Livello

Università

Esercizi

10, con soluzione ragionata

Prerequisiti

Calcolo combinatorio, probabilità di base

FORMULE DI RIFERIMENTO

VALORE ATTESO

$$E(X) = \sum_i x_i P(x_i)$$

PERMUTAZIONI SEMPLICI

$$P(n, r) = \frac{n!}{(n-r)!}$$

DISTRIBUZIONE BINOMIALE

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

VARIANZA

$$\text{Var}(X) = \sum_i x_i^2 P(x_i) - [E(X)]^2$$

COMBINAZIONI SEMPLICI

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

MEDIA E VARIANZA BINOMIALE

$$E(X) = np \quad \text{Var}(X) = np(1-p)$$

ESERCIZIO 1

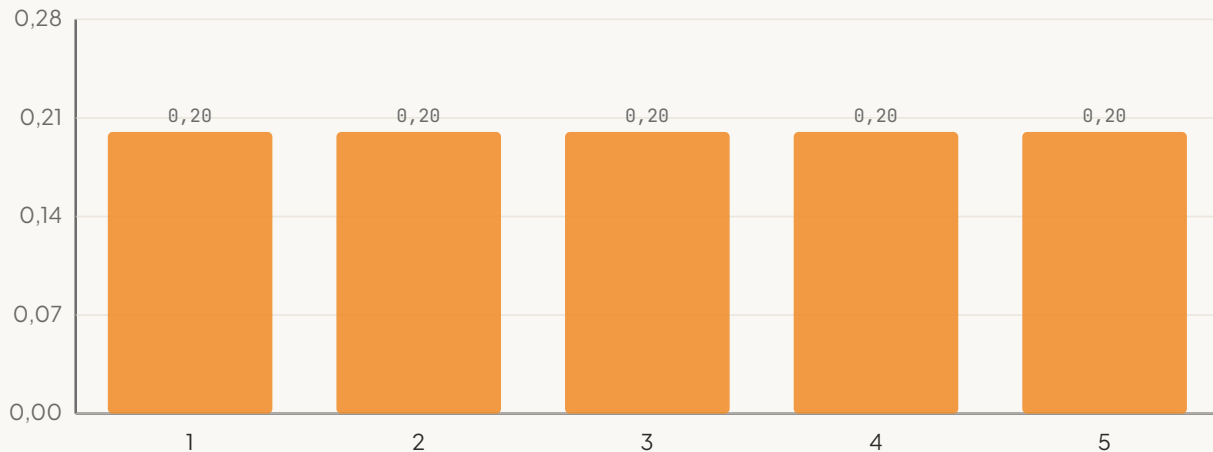
Distribuzione uniforme discreta: rappresentazione, media e varianza

Rappresenta graficamente la distribuzione uniforme di probabilità descritta dalla funzione

$$P(x_i) = \frac{1}{5}, \quad i = 1, 2, 3, 4, 5$$

Calcola inoltre media e varianza di questa distribuzione di probabilità.

SOLUZIONE



Distribuzione uniforme su {1, 2, 3, 4, 5} — ogni valore ha probabilità $1/5 = 0,20$

$$E(X) = \sum_{i=1}^5 x_i \cdot \frac{1}{5} = \frac{1 + 2 + 3 + 4 + 5}{5} = 3$$

$$\text{Var}(X) = \sum_{i=1}^5 x_i^2 \cdot \frac{1}{5} - 3^2 = 11 - 9 = 2$$

MEDIA

$$E(X) = 3$$

VARIANZA

$$\text{Var}(X) = 2$$

OSSERVAZIONE

Per una distribuzione uniforme discreta su n valori consecutivi la media coincide sempre con il valore centrale della serie: qui infatti 3 è il valore di mezzo tra 1 e 5. È un buon controllo rapido del risultato.



ESERCIZIO 2

Permutazioni: l'ordine di arrivo di una gara

Otto velocisti competono in una gara di 100 metri.

A) Quanti sono i possibili ordini di arrivo degli 8 velocisti?

B) 3 velocisti otterranno una medaglia. Quante sono le triplete possibili (ordinate) dei medagliati?

SOLUZIONE

A) Ogni diverso ordine di arrivo è una permutazione semplice degli 8 velocisti:

$$8! = 40\,320$$

B) L'ordine sul podio conta (oro, argento, bronzo), quindi si tratta di una disposizione semplice di 8 elementi in classi di 3:

$$P(8, 3) = \frac{8!}{(8-3)!} = 8 \cdot 7 \cdot 6 = 336$$

A

40.320 ordini

B

336 podi possibili

ESERCIZIO 3

Permutazioni e combinazioni: la gara di galoppo

Sei cavalli competono in una gara di galoppo.

A) In quanti modi diversi possono ordinarsi al traguardo i 6 cavalli?

B) Se scommettessi su una particolare coppia di cavalli, arrivo indifferentemente 1° e 2° (l'ordine non conta), quante sarebbero in tutto le coppie di cavalli da considerare prima di decidere su chi scommettere?

SOLUZIONE

A) Qui l'ordine di arrivo conta (come nell'Esercizio 2): è una permutazione semplice di 6 elementi.

$$6! = 720$$

B) La coppia di podio non è ordinata (non importa chi dei due arriva 1° e chi 2°, conta solo quali due cavalli): è una combinazione semplice di 6 elementi in classi di 2.

$$\binom{6}{2} = \frac{6!}{2! 4!} = 15$$

A

720 ordini

B

15 coppie possibili

COMBINAZIONI VS. PERMUTAZIONI

Attenzione a non confondere le due situazioni: se la domanda A chiedesse invece "quante sono le possibili *combinazioni* di arrivo dei 6 cavalli" (cioè quanti sottoinsiemi di 6 cavalli scelti fra 6, senza ordine), la risposta

sarebbe banale: $\binom{6}{6} = \frac{6!}{6! 0!} = 1$ — un solo modo di scegliere "tutti e 6", perché l'ordine viene ignorato. È

proprio l'ordine di arrivo a rendere il problema interessante, quindi il calcolo corretto è la permutazione (punto A sopra), non la combinazione.

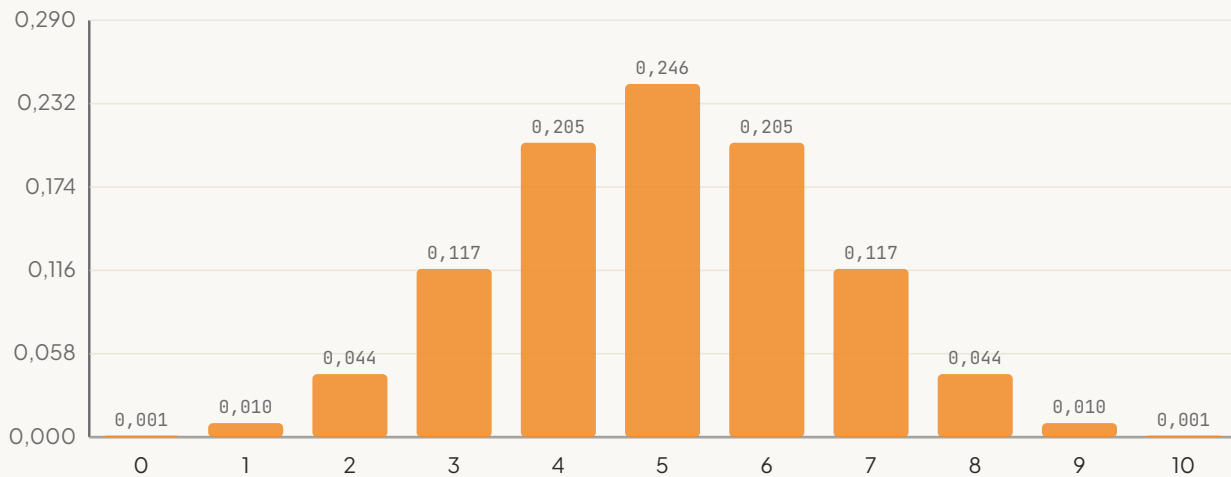
ESERCIZIO 4

Distribuzione binomiale: grafico, media e deviazione standard

Produci il grafico di una distribuzione binomiale con $p = 0,5$ e $n = 10$. Determina media e deviazione standard della distribuzione.

SOLUZIONE

$$P(X = k) = \binom{10}{k} \left(\frac{1}{2}\right)^{10}$$



Distribuzione binomiale $\text{Bin}(n = 10, p = 0,5)$ — massimo a $k = 5$, simmetrica

K (SUCCESSI)	P(X = K)
0	0,000977
1	0,009766
2	0,043945
3	0,117188
4	0,205078
5	0,246094
6	0,205078
7	0,117187
8	0,043945
9	0,009766
10	0,000977

$$E(X) = np = 10 \cdot 0.5 = 5$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{2.5} \approx 1.581$$

MEDIA

$$E(X) = 5$$

DEV. STANDARD

$$\sigma \approx 1,581$$



ESERCIZIO 5

Test d'ipotesi con la distribuzione binomiale

Una ricerca è stata effettuata per valutare se un particolare tipo di deodorante alla fragola fosse più piacevole di uno alla menta. 10 persone sono state testate e 9 hanno scelto il deodorante alla fragola.

A) Usa la distribuzione binomiale relativa all'Esercizio 4 per decidere se l'ipotesi nulla di equiprobabilità di scelta dei due deodoranti va rigettata.

B) Se 8 persone avessero scelto il deodorante alla fragola, avresti rigettato l'ipotesi nulla?

SOLUZIONE

Ipotesi nulla H_0 : $p = 0,5$ (i due deodoranti sono scelti con la stessa probabilità). Sotto H_0 , $X \sim \text{Bin}(10; 0,5)$ come nell'Esercizio 4. Test a due code, soglia $\alpha = 0,05$.

A) Con 9 successi su 10, la probabilità (a due code) di osservare un risultato altrettanto o più estremo è:

$$2 \cdot P(X \geq 9) = 2 \left[\binom{10}{9} (0.5)^{10} + \binom{10}{10} (0.5)^{10} \right] \approx 0.0215$$

CONCLUSIONE A

$0,0215 < 0,05 \rightarrow$ si rigetta H_0

La preferenza per il deodorante alla fragola è statisticamente significativa.

B) Con 8 successi su 10:

$$2 \cdot P(X \geq 8) \approx 0.1094$$

CONCLUSIONE B

$0,1094 > 0,05 \rightarrow$ non si rigetta H_0

Con un solo caso in meno, il risultato non è più sufficiente per concludere una preferenza significativa.

PERCHÉ "A DUE CODE"

Il test è bilaterale perché l'ipotesi nulla afferma semplicemente "nessuna preferenza", senza specificare a priori quale dei due deodoranti sarebbe stato preferito: un risultato estremo in favore della fragola conta quanto uno, altrettanto estremo, in favore della menta.

ESERCIZIO 6

Individuare una moneta truccata

Una moneta viene lanciata 16 volte e devi decidere, sulla base del risultato ottenuto (il numero di teste osservate), se la moneta è truccata. Come puoi decidere se la moneta è truccata utilizzando i dati ottenuti da un esperimento come quello descritto?

SOLUZIONE

Una moneta non truccata lanciata 16 volte può dare da 0 a 16 teste. Sotto l'ipotesi H_0 "moneta onesta" ($p = 0,5$), il numero di teste X segue:

$$P(X = k) = \binom{16}{k} \left(\frac{1}{2}\right)^{16}$$

K	P(X=K)
0	0,000015
1	0,000244
2	0,001831
3	0,008545
4	0,027771
5	0,066650
6	0,122192
7	0,174561
8	0,196381

K	P(X=K)
9	0,174561
10	0,122192
11	0,066650
12	0,027771
13	0,008545
14	0,001831
15	0,000244
16	0,000015

Si cerca la regione di rigetto più ampia con probabilità totale (a due code) non superiore a 0,05:

FINO A $K = 3$ (E SIMMETRICO 13-16)

$$2 \cdot P(X \leq 3) = 2 \cdot 0.01064 \approx 0.0213$$

Sotto soglia: regione accettabile.

INCLUDENDO $K = 4$ (E 12)

$$2 \cdot P(X \leq 4) = 2 \cdot 0.03841 \approx 0.0768$$

Supera 0,05: non accettabile.

REGOLA DI DECISIONE

Rigetta H_0 se $X \in \{0, 1, 2, 3\} \cup \{13, 14, 15, 16\}$

In tal caso la probabilità totale di rigettare erroneamente una moneta onesta è $2 \times 0,01064 \approx 0,0213$, sotto la soglia del 5%. Aggiungere anche 4 e 12 farebbe salire l'errore complessivo a $\approx 7,7\%$, troppo alto.

ESERCIZIO 7

Ammissione a un'accademia: probabilità su un gruppo

La probabilità che uno studente passi al primo tentativo l'esame di ammissione all'accademia aereospaziale di Forlimpopoli è del 20%.

A) Qual è la probabilità che almeno uno studente, in un gruppo di 10, passi il test al primo tentativo?

B) E qual è la probabilità che esattamente 7 studenti su 10 passino il test d'ammissione al primo tentativo?

SOLUZIONE

Sia X il numero di studenti (su 10) che passano il test al primo tentativo. Con probabilità individuale $p = 0,2$, $X \sim \text{Bin}(10; 0,2)$:

$$P(X = k) = \binom{10}{k} (0.2)^k (0.8)^{10-k}$$

K (STUDENTI SU 10)	P(X = K)
0	0,107374
1	0,268435
2	0,301990
3	0,201327
4	0,088080
5	0,026424
6	0,005505
7	0,000786
8	0,000074
9	0,000004
10	0,000000

A) "Almeno uno" è il complementare di "nessuno passa" ($X = 0$):

$$P(X \geq 1) = 1 - (0.8)^{10} \approx 0.893$$

A

P(X ≥ 1) ≈ 0,893

C'è una probabilità dell'89,3% che almeno uno dei 10 studenti passi al primo tentativo.

B) "Esattamente 7 su 10" è il singolo valore $P(X = 7)$:

$$P(X = 7) = \binom{10}{7} (0.2)^7 (0.8)^3 \approx 0.001$$

B

P(X = 7) ≈ 0,001

Con $p = 20\%$, avere 7 successi su 10 è un evento molto raro (0,08%) — coerente con il fatto che il valore atteso è $np = 2$.

ESERCIZIO 8

Confronto fra due test: matematica vs. grammatica

Un campione di 20 bambini è testato in un esercizio di matematica e in uno di grammatica. 14 bambini risultano superare il test di matematica con un punteggio più alto di quello di grammatica (i voti in entrambi i test sono su scale standardizzate con media e deviazione standard comuni). Possiamo concludere che la popolazione da cui i bambini sono stati estratti è più brava in matematica che in grammatica?

SOLUZIONE

Si tratta di un **test del segno**: sotto l'ipotesi nulla H_0 "nessuna differenza sistematica" ($p = 0,5$ che un bambino sia migliore in matematica), il numero X di bambini con matematica > grammatica segue $\text{Bin}(20; 0,5)$. Si osserva $X = 14$.

$$P(X \geq 14) = \sum_{k=14}^{20} \binom{20}{k} (0.5)^{20} \approx 0.0577$$

Test a due code (l'ipotesi nulla non specifica a priori una direzione):

$$2 \cdot P(X \geq 14) \approx 0.1153$$

CONCLUSIONE

0,058 (una coda) / 0,115 (due code) > 0,05 → non si rigetta H_0

Il risultato è vicino alla soglia di significatività ma non la supera: **non** abbiamo prove statisticamente sufficienti, al livello del 5%, per affermare che la popolazione sia sistematicamente più brava in matematica che in grammatica. Il campione mostra una tendenza in quella direzione, ma non abbastanza forte da escludere il caso.

NOTA METODOLOGICA

Con 14 "successi" su 20 il risultato è al limite: una sola osservazione in più (15 su 20, $p \approx 0,021$) avrebbe portato a rigettare H_0 . Il test del segno usa solo il segno della differenza fra le due misure, non la loro grandezza — è un test semplice ma con potere statistico limitato, come si vede meglio negli Esercizi 9 e 10.

ESERCIZIO 9

Test del segno: l'effetto di una dieta (n = 10)

Un campione di 10 donne viene pesato (in kg) prima e dopo un periodo di dieta di 2 settimane. Usa il test del segno per decidere se la dieta induce l'effetto desiderato (perdita di peso).

SOLUZIONE

#	PRIMA	DOPO	Δ	#	PRIMA	DOPO	Δ
1	75	72	-3	6	79	68	-11
2	76	65	-11	7	69	72	+3
3	65	56	-9	8	76	68	-8
4	77	79	+2	9	72	65	-7
5	67	66	-1	10	67	64	-3

8 donne su 10 hanno perso peso (segno atteso), 2 lo hanno preso. Sotto H_0 : $p = 0,5$, con $X \sim \text{Bin}(10; 0,5)$:

$$P(X \geq 8) = \sum_{k=8}^{10} \binom{10}{k} (0.5)^{10} \approx 0.0547$$

CONCLUSIONE

0,0547 > 0,05 → non si rigetta H_0

Anche a una coda il risultato resta (di pochissimo) sopra la soglia del 5%: sulla base di questi 10 dati non si può concludere che la dieta abbia avuto un effetto statisticamente significativo, pur essendo il risultato molto vicino al limite.

ESERCIZIO 10

Test del segno: la stessa dieta con un campione più ampio (n = 20)

Un campione di 20 donne viene pesato (in kg) prima e dopo lo stesso periodo di dieta di 2 settimane (le prime 10 sono quelle dell'Esercizio 9, più altre 10 nuove osservazioni).

A) Usa il test del segno per decidere se la dieta induce l'effetto desiderato.

B) Confrontando i risultati di questo esercizio con quelli dell'Esercizio 9, cosa puoi inferire?

SOLUZIONE

#	PRIMA	DOPO	Δ	#	PRIMA	DOPO	Δ
1	75	72	-3	11	65	56	-9
2	76	65	-11	12	67	64	-3
3	65	56	-9	13	76	65	-11
4	77	79	+2	14	72	65	-7
5	67	66	-1	15	67	66	-1
6	79	68	-11	16	79	68	-11
7	69	72	+3	17	76	68	-8
8	76	68	-8	18	75	72	-3
9	72	65	-7	19	69	72	+3
10	67	64	-3	20	77	79	+2

A) 16 donne su 20 hanno perso peso. Sotto H_0 : $p = 0,5$, con $X \sim \text{Bin}(20; 0,5)$:

$$P(X \geq 16) = \sum_{k=16}^{20} \binom{20}{k} (0.5)^{20} \approx 0.0059$$

$$2 \cdot P(X \geq 16) \approx 0.0118$$

CONCLUSIONE A

0,0059 (una coda) / 0,0118 (due code) < 0,05 → si rigetta H_0

Con il campione più ampio la dieta risulta avere un effetto statisticamente significativo.

B) CONFRONTO CON L'ESERCIZIO 9

Il test del segno tiene conto solo del segno della differenza, non della sua grandezza: la proporzione di "successi" è la stessa (80%) in entrambi gli esercizi, eppure solo con $n = 20$ il risultato diventa significativo. Un campione più numeroso aumenta il **potere statistico** del test — la capacità di rilevare un effetto reale quando c'è — anche a parità di proporzione osservata.