

Esercitazioni di statistica

Variabili casuali discrete

Stefania Spina

Università di Napoli Federico II

stefania.spina@unina.it

19 Novembre 2014

Le variabili casuali: Definizioni

Le **variabili casuali** agevolano l'interpretazione dei risultati di un esperimento casuale definendo una corrispondenza tra eventi e numeri reali.

La **variabile** è detta **casuale** perchè il valore numerico che assume dipende dall'esito dell'esperimento.

Una **variabile casuale** è una funzione che associa agli eventi dello spazio campionario S un numero reale.

La variabile casuale è **discreta** se assume valori in un insieme discreto (es. *il numero di pezzi difettosi in una produzione, il numero di studenti laureati, il numero di errori in un testo,...*)

La variabile casuale è **continua** se assume valori in un insieme continuo, in un intervallo limitato o illimitato (es. *il peso di una confezione di biscotti, il reddito medio annuo di una famiglia, la durata di una batteria,...*)

Le variabili casuali: Definizioni

Possiamo distinguere:

V.c. discrete	V.c. continue
Bernoulli	Normale
Binomiale	Lognormale
Poisson	Gamma
Ipergeometrica	Chi-quadrato
Uniforme	t di Student
	F di Fisher-Snedecor
	Esponenziale negativa
	Uniforme

Iniziamo con le variabili casuali discrete

Esercizio 1

Un commerciante fa un'ordinazione di 100 transistor.

La sua politica consiste nel rifiutare la partita se, su un campione casuale di 7, più di uno risulta difettoso.

Supponendo che tra i 100 transistor ordinati 4 siano difettosi, **qual è la probabilità che il commerciante rifiuti la partita?**

Esercizio 1

Evento “successo”: il pezzo è difettoso

$$n = 7$$

$$p = \frac{4}{100} = 0.04$$

Variabile casuale (X): *numero di pezzi difettosi su un totale di 7*

$$E(X) = n * p = 7 * 0.04 = 0.28$$

$$Var(X) = np * (1 - p) = 7 * 0.04 * 0.96 = 0.269 \Rightarrow \sigma = \sqrt{Var(X)} = \sqrt{0.269} = 0.518$$

Esercizio 1

$$\begin{aligned}P(X > 1) &= [P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)] = \\&= 1 - [P(X = 0) + P(X = 1)]\end{aligned}$$

$$\text{dove } P(X = 0) = 0.96^7 = 0.751$$

$$\text{dove } P(X = 1) = \binom{n}{x} p^x (1 - p)^{n-x} = \frac{7!}{1! * 6!} 0.04^1 (1 - 0.04)^{7-1} = 0.219$$

$$P(X > 1) = 1 - [P(X = 0) + P(X = 1)] = 1 - [0.751 + 0.219] = 1 - 0.970 = 0.03$$

Esercizio 2

Si ipotizzi che la prova di esame consista in 20 domande a risposta multipla, ciascuna con quattro possibili risposte di cui una sola esatta, e che la prova scritta si consideri superata se si risponde esattamente ad almeno 18 domande.

Supponendo che uno studente sia sicuro di 15 risposte e tiri a indovinare per le altre 5, qual è la probabilità che passi la prova?

Esercizio 2

Evento “successo”: la risposta è esatta

$$n = 5$$

$$p = \frac{1}{4} = 0.25$$

Variabile casuale (X): *numero di risposte esatte su un totale di 5 domande*

$$E(X) = n * p = 5 * 0.25 = 1.25$$

$$Var(X) = np * (1 - p) = 5 * 0.25 * 0.75 = 0.938 \Rightarrow \sigma = \sqrt{Var(X)} = \sqrt{0.938} = 0.969$$

Esercizio 2

$$P(X \geq 3) = [P(X = 3) + P(X = 4) + P(X = 5)]$$

$$P(X = 3) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{5!}{3! * 2!} 0,25^3 (1-0.25)^{5-3} = 10 * 0.0156 * 0.5625 = 0.087$$

$$P(X = 4) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{5!}{4! * 1!} 0.25^4 (1-0.25)^{5-4} = 5 * 0.0039 * 0.75 = 0.0146$$

$$P(X = 5) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{5!}{5! * 0!} 0.25^5 (1-0.25)^{5-5} = 1 * 0.000977 * 1 = 0.000977$$

$$P(X \geq 3) = [P(X = 3) + P(X = 4) + P(X = 5)] = 0.087 + 0.0146 + 0.0010 = 0.1026$$

Esercizio 3

Si consideri l'esperimento consistente nel lancio di una moneta 120 volte. Trovare la probabilità che esca testa:

- tra il 40% e il 60% dei lanci;
- più dei $\frac{5}{8}$ del totale dei lanci.

Ci sono due metodi per risolvere il primo punto.

Esercizio 3

- tra il 40% e il 60% dei lanci

Primo metodo

La v.c. numero di teste in 120 lanci si distribuisce come una binomiale con parametri ($n = 120$, $p = \frac{1}{2}$). Il numero di successi corrispondenti al 40% totale di 120 prove è 48, mentre il 60% corrisponde a 72. di

I parametri della normale cui si approssima la binomiale sono:

$$\mu = np = 120 * \frac{1}{2} = 60$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{120 * \frac{1}{2} * \frac{1}{2}} = 5.48$$

Esercizio 3

Standardizzando i valori e considerando la correzione per la continuità si avrà:

$$z_1 = \frac{47.5 - 60}{5.48} = -2.28$$

$$z_2 = \frac{72.5 - 60}{5.48} = 2.28$$

Dalla simmetria della normale segue che:

$$P(-2.28 \leq z \leq 2.28) = 2 * P(0 \leq z \leq 2.28) = 2 * 0.4887 = 0.97$$

Esercizio 3

Secondo metodo

Il secondo metodo che permette di risolvere il primo punto consiste nel ricorrere alla distribuzione delle proporzioni campionarie, la cui

media è data da $\mu_p = p = 0.5$ e $\sigma_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.25}{120}} = 0.0456$

Standardizzando il problema rispetto ai parametri della distribuzione campionaria

$$z_1 = \frac{0.4 - 0.5}{0.0456} = -2.19$$

$$z_2 = \frac{0.6 - 0.5}{0.0456} = 2.19$$

Dalla simmetria della normale segue che:

$$P(-2.19 \leq z \leq 2.19) = 2 * P(0 \leq z \leq 2.19) = 2 * 0.4857 = 0.97$$

Esercizio 3

Il risultato non coincide con quanto ottenuto utilizzando l'approssimazione della binomiale alla normale perchè la v.c. proporzione è una variabile discreta.

Per tenere conto di questo bisogna sottrarre $\frac{1}{2N} = \frac{1}{2 \cdot 120} = 0.00417$ a 0.4 ed aggiungere la stessa quantità a 0.6.

Pertanto gli estremi dell'intervallo di valori standardizzati è:

$$z_1 = \frac{0.4 - 0.00417 - 0.5}{0.0456} = -2.28$$

$$z_2 = \frac{0.6 + 0.00417 - 0.5}{0.0456} = 2.28$$

Che porta alla coincidenza tra i risultati nei due metodi.

Esercizio 3

- più dei $\frac{5}{8}$ del totale dei lanci.

La proporzione corrispondente a $\frac{5}{8} = 0.6250$. Standardizzando il problema tenendo conto del fatto che la proporzione è una variabile discreta si ottiene

$$z = \frac{0.6250 - 0.00417 - 0.5}{0.0456} = 2.65$$

La probabilità corrispondente è

$$P(z \geq 2.65) = 1 - P(z < 2.65) = 1 - 0.996 = 0.004$$

Esercizio 4

Un componente elettronico è soggetto a guasti casuali che seguono un processo di Poisson. In 10 settimane mediamente si verifica un guasto. Si determinino:

- 1 il numero medio di guasti in un anno;
- 2 le probabilità che non ci siano guasti per un anno;
- 3 le probabilità che non ci sia più di un guasto in 20 settimane.

Esercizio 4

La v.c. di Poisson è una variabile discreta e viene, di solito, utilizzata per **analizzare fenomeni connessi a conteggi**: n di automobili che passano in un dato punto in un determinato intervallo di tempo, n di clienti serviti in una fila in un determinato tempo, arrivo di aerei in un aeroporto, arrivi al pronto soccorso di un ospedale, chiamate al 113, numero di incidenti che avvengono in una città in un **determinato intervallo di tempo**, ecc...



E' derivata, quindi, come il numero di X eventi che si verificano in un definito intervallo di tempo e che può assumere valori $x = 0, 1, 2, \dots$

Esercizio 4

Un componente elettronico è soggetto a guasti casuali che seguono un processo di Poisson. In 10 settimane mediamente si verifica un guasto. Si determinino:

1) il numero medio di guasti in un anno.

$$\lambda_{10s} = 1; 1 \text{ anno} = 52s \rightarrow \lambda_{1s} = \frac{1}{10} = 0.1 \rightarrow \lambda_{1a} = 0.1 * 52 = 5.2$$

Si calcolino, inoltre, le probabilità che:

2) non ci siano guasti per un anno;

$$Pr(X = x) = \exp(-\lambda) \frac{\lambda^x}{x!} \text{ per } x = 0, 1, 2, \dots$$

$$Pr(X = 0) = \exp(-5.2) \frac{\lambda^0}{0!} = 0.0055$$

3) non ci sia più di un guasto in 20 settimane.

$$\lambda_{20s} = 1 * 2 = 2$$

$$\begin{aligned} Pr(X < 1) &= \sum_{x=0}^1 \exp(-\lambda) \frac{\lambda^x}{x!} = \exp(-2) \frac{\lambda^0}{0!} + \exp(-2) \frac{\lambda^1}{1!} = \\ &= \exp(-2) + 2\exp(-2) = 3\exp(-2) \cong 0,406 \end{aligned}$$

Esercizio 5

Il numero ideale di studenti per il primo anno di un corso in una certa Facoltà è di 150.

La Facoltà, sapendo dall'esperienza passata che solo il 30% degli studenti ammessi segue le lezioni, adotta la politica di accettare le iscrizioni di 450 studenti.

Si calcoli la probabilità che più di 150 studenti del primo anno frequentino le lezioni.

Esercizio 5

Si definisca la v.c. X come il numero di studenti che frequentano, ciascuno studente iscritto corrisponde ad una prova Bernoulliana il cui esito può essere frequenta o non frequenta.

X si distribuisce pertanto secondo una distribuzione binomiale di parametri $n = 450$ e la probabilità di successo (lo studente iscritto frequenta) è $p = 0.3$.

Per il teorema del limite centrale, poichè X è la somma di n v.c. Bernoulliana X_i , ciascuna con media $E[X_i] = p$ e varianza pari a $\text{var}(X_i) = p(1 - p)$, allora X si distribuisce approssimativamente come una normale con media $\mu = np$

X = numero di studenti che frequentano $\rightarrow \text{Bin} \sim (450, 0.3)$

$$P(X > 150) = \sum_{i=151}^{\infty} P(X = i)$$

Esercizio 5

Approssimazione normale della v.c. Binomiale: La probabilità cercata è $P(X > 150.5)$ (lo '+ 0.5' in aggiunta è dovuto alla correzione di continuità).

$$P(X > 150) = \sum_{i=151}^{\infty} P(i - 0.5 < X < i + 0.5) = P(X > 150.5)$$

$$P(X > 150.5) = P\left(\frac{X - \mu}{\sigma} > \frac{150.5 - 450 * 0.3}{\sqrt{450 * 0.3 * 0.7}}\right) = P(z > 1.59) \cong 0.06$$

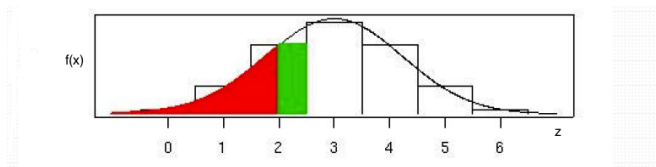
solo il 6% circa degli studenti che decidono di seguire superano il numero raccomandato di 150

Esercizio 5

Funzioni di massa Binomiali che convergono ad una densità Normale.

Correzione per la continuità'!

Supponiamo di avere una v.c. $X \sim B(6, 0.5)$ e vogliamo calcolare la $P(X \leq 2)$:



Calcolando esattamente la $P(X \leq 2)$ non si includerebbe l'area verde, ovvero la probabilità calcolata sarebbe inferiore a quella effettiva.



Correzione per la continuità: aumentiamo di 0.5 (o diminuiamo in caso di $P(X \geq 2)$)

$$P(X \leq 2) \rightarrow P(X \leq 2.5)$$