Foglio di esercizi 6

Discussione soluzioni: 30.04.2022

1. Riempire la tabella seguente per X, Y indipendenti (in ogni riga si scrivano, se possibile, solo funzioni semplici dei parametri corrispondenti a quella riga)

Distribuzione di X, Y	Parametri	Media	Varianza	Distribuzione di $X + Y$
Bernoulli(n, p)	p	p	$p(1-p)$	
Uniforme$(\{a, \ldots, b\})$	(a, b)			
Binomiale(n, p)	(n, p)			
Geometrica(p)	p			
Poisson(λ)	λ			
Normale(μ, σ^2)	(μ, σ^2)	μ	σ^2	$\mathcal{N}(\mu + \mu, \sigma^2 + \sigma^2)$

2. Si avvicina una grande elezione in Regione Immaginaria, una regione di $100'000$ persone. In Regione Immaginaria, il 51% delle persone preferiscono il candidato A e il 49% il candidato B.
   (a) Un sondaggio sceglie a caso (con ricollocamento) 100 persone dalla popolazione. Scrivere una formula per calcolare *esattamente* la probabilità che almeno il 51% delle persaone intervistate preferiscano il candidato A (non si valuti questa espressione numericamente).
   (b) Scrivere una formula per calcolare la probabilità dell'evento descritto al punto precedente quando la scelta di persone per il sondaggio viene effettuata *senza ricollocamento* (anche in questo caso non si valuti questa espressione numericamente).
   (c) Si usi il Teorema del Limite Centrale per approssimare la probabilità dell'evento descritto al punto (a).
   (d) Quante persone si dovrebbe intervistare (con ricollocamento) durante un sondaggio per essere sicuri al 99% che la proporzione osservata di persone che preferiscono il candidato A sia entro l'1% dal suo vero valore? (Si usi il Teorema del Limite Centrale per approssimare la risposta)

3. Una fabbrica produce delle lampadine speciali. In media, il 2.3% delle lampadine prodotte ha un difetto di produzione. Un controllo qualità prende 1000 lampadine, scelte a caso tra quelle prodotte (quindi assunte indipendenti), e le testa per vedere se sono difettose. Usando il teorema del limite centrale, calcolare la probabilita che nel campione ci siano più di 28 lampadine difettose. Fare questo calcolo con e senza correzione di continuità.

4. Sia $(X_n)_{n \geq 0}$ una successione di variabili aleatorie indipendenti e identicamente distribuite con valore atteso $\mathbb{E}(X_i) = \mu = 9$ e $\mathrm{SD}(X_i) = \sigma = 3$. Poniamo $\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$.
   (a) Approssimativamente, quanto ci aspettiamo che valga $\bar{X}_n$ per n molto grande? Motivare la propria risposta con un risultato visto in classe.
   (b) Stimare con la disuguaglianza di Chebyshev il valore di
   $$\mathbb{P}[|\bar{X}_n - 9| \geq 30]$$

5. Circa il 5% dei passeggeri con dei biglietti perde l'aereo. Su un volo con 200 posti, assumiamo che una compagnia aerea accetti 212 riservazioni. Assumendo che ogni passeggero arrivi in tempo all'imbarco indipendentemente da tutti gli altri passeggeri, si usi (se possibile) l'approssimazione di Poisson per stimare la probabilità che il volo sia overbooked.

Distribuzione di X, Y	Parametri	Media	Varianza	Distribuzione di $X+Y$
Bernoulli($1, p$)	p	p	$p(1-p)$	$Bin(2,p)$
Uniforme($\{a, \dots, b\}$)	(a,b)	$\frac{a+b}{2}$	$(b-a+1)^2-1/12$	—
Binomiale(n,p)	(n,p)	$n \cdot p$	$n \cdot p(1-p)$	$Bin(2n,p)$
Geometrica(p)	p	$1/p$	$(1-p)/p^2$	—
Poisson(λ)	λ	λ	λ	$Pois(\lambda+\lambda)$
Normale(μ, σ^2)	(μ, σ^2)	μ	σ^2	$\mathcal{N}(\mu+\mu, \sigma^2+\sigma^2)$

Esponenziale

$X \sim Bin(n,p) \qquad X = \mathbb{1}_1 + \mathbb{1}_2 \cdots + \mathbb{1}_n$

$$\mathbb{1}_j = \begin{cases} 1 & \text{se successo all'esperimento } j \\ 0 & \text{altrimenti} \end{cases}$$

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{1}_1 + \dots + \mathbb{1}_n) = \mathbb{E}(\mathbb{1}_1) + \dots + \mathbb{E}(\mathbb{1}_n) = np$$

$$Var(X) = Var(\mathbb{1}_1 + \dots \mathbb{1}_n) = Var(\mathbb{1}_1) + \dots + Var(\mathbb{1}_n) = np(1-p)$$

$Y \sim Unif(\{a, \dots, b\}) = a + Z \qquad Z \sim Unif(\{0, 1, \dots, b-a\})$

$$\mathbb{E}(Y) = \mathbb{E}(a+Z) = a + \mathbb{E}(Z) = a + \sum_{j=0}^{b-a} j \frac{1}{b-a+1}$$

$$= a + \frac{1}{b-a+1} \cdot \sum_{j=1}^{b-a} j = a + \frac{(b-a)\cancel{(b-a+1)}}{2} \cdot \frac{1}{\cancel{b-a+1}}$$

$$= \frac{2a}{2} + \frac{b-a}{2} = \frac{a+b}{2}$$

$$Var(Y) = Var(a+Z) = Var(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2$$

$$= \sum_{j=0}^{b-a} \frac{1}{b-a+1} j^2 - \left(\frac{a+b}{2}\right)^2 = \dots = \frac{(b-a+1)^2 - 1}{12}$$

2. Si avvicina una grande elezione in Regione Immaginaria, una regione di $100'000$ persone. In Regione Immaginaria, il 51% delle persone preferiscono il candidato A e il 49% il candidato B.

   (a) Un sondaggio sceglie a caso (con ricollocamento) 100 persone dalla popolazione. Scrivere una formula per calcolare *esattamente* la probabilità che almeno il 51% delle persaone intervistate preferiscano il candidato A (non si valuti questa espressione numericamente).

   (b) Scrivere una formula per calcolare la probabilità dell'evento descritto al punto precedente quando la scelta di persone per il sondaggio viene effettuata *senza ricollocamento* (anche in questo caso non si valuti questa espressione numericamente).

   (c) Si usi il Teorema del Limite Centrale per approssimare la probabilità dell'evento descritto al punto (a).

   (d) Quante persone si dovrebbe intervistare (con ricollocamento) durante un sondaggio per essere sicuri al 99% che la proporzione osservata di persone che preferiscono il candidato A sia entro l'1% dal suo vero valore? (Si usi il Teorema del Limite Centrale per approssimare la risposta)

a) $\mathbb{1}_j = \begin{cases} 1 & \text{se la } j\text{-esima persona preferisce } A \\ 0 & \text{altrimenti} \end{cases}$

$$\mathbb{P}(\mathbb{1}_j = 1) = p = 0.51$$

$$X = \mathbb{1}_1 + \mathbb{1}_2 + \dots + \mathbb{1}_{100} = S_n \sim Bin(100, p)$$

$$\mathbb{P}(X \geq 51) = \sum_{k=51}^{100} \mathbb{P}(X = k) = \sum_{k=51}^{100} \binom{100}{k} 0.51^k (1-0.51)^{100-k}$$

b) Senza ricollocamento $\quad N_A =$ persone in totale che preferiscono A

$= 100000 \cdot 0.51 = 51000.$

$$\mathbb{P}(\mathbb{1}_1 = 1, \dots, \mathbb{1}_{51} = 1, \mathbb{1}_{52} = 0, \dots \mathbb{1}_{100} = 0) =$$

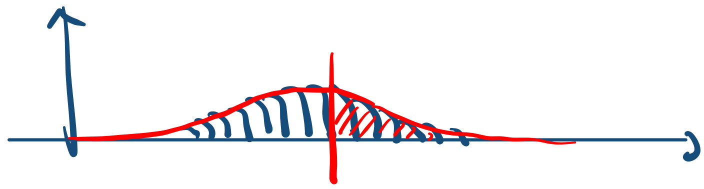
$$= \frac{51000 \cdot 50999 \cdots 50950 \cdot 49000 \cdot 48999 \cdots 48952}{100000 \cdot 99999 \; 99998 \cdots 99901}$$

$$= \bar{P}_{51}$$

$$\mathbb{P}(X=51) = \binom{100}{51}\cdot \bar{p}_{51}$$

$$\Longrightarrow \mathbb{P}(X\geq 51) = \sum_{k=51}^{100} \mathbb{P}(X=k) \qquad (\text{difficile}).$$

c) $X \approx \underline{N(\mu_\#, \sigma_\#^2)}$



$$\mathbb{P}(X\geq 51) = \int_{51}^{\infty} f_{\mu_\#,\sigma_\#^2}(x)\,dx =$$

$$= \Phi\left(\frac{\infty-\mu_\#}{\sigma_\#}\right) - \Phi\left(\frac{51-\mu_\#}{\sigma_\#}\right) =$$

$$= 1 - \Phi\left(\frac{51-\mu_\#}{\sigma_\#}\right) = 1 - \Phi\left(\frac{51-51}{\sigma_\#}\right)$$

$$= 1 - \Phi(0) = 0.5$$

$$\begin{cases} \mu_\# = n\cdot p \\ \sigma_\#^2 = np(1-p) \end{cases} \qquad \begin{cases} \mu_\# = 51 \\ \sigma_\# = \sqrt{51\cdot(0.49)} \end{cases}$$

Con correzione di continuità

$$\mathbb{P}(X\geq 51) = 1 - \mathbb{P}(X<51) = 1 - \mathbb{P}(X\leq 50)$$

$$= 1 - \Phi\left(\frac{50+\frac{1}{2}-51}{\sqrt{51\cdot 0.49}}\right) = 1 - \Phi\left(-\frac{1}{2\sqrt{51\cdot 0.49}}\right)$$

$$= \dots$$

d) $\mathbb{P}\left(\bar{X}_n \in \left(0.51 - \underbrace{0.01}_{\gamma}, 0.51 + 0.01\right)\right) = 0.99$

$$\bar{X}_n = \frac{1}{n} X$$

Identifichiamo:

- $\gamma = 0.01$
- $z^* = 2.57$
- $\sigma_n^2 = Var(\bar{X}_n) = Var\left(\frac{1}{n} \cdot X\right)$

$$= \frac{1}{n^2} Var(X) = \frac{\cancel{n} p(1-p)}{n^{\cancel{2}}}$$

$$= \frac{p(1-p)}{n}$$

$$0.01 = 2.57 \cdot \sqrt{\frac{0.51(0.49)}{n}}$$

$\xrightarrow[\text{per } n]{\text{risolvo}}$ $n = 2250$

Riempire la tabella seguente per X, Y indipendenti (in ogni riga si scrivano, se possibile, solo funzioni semplici dei parametri corrispondenti a quella riga)

Distribuzione di X, Y	Parametri	Media	Varianza	Distribuzione di $X + Y$
Bernoulli(n, p)	p	p	$p(1-p)$	$Bin(2,p)$ se $X \sim Ber(p)$ $Y \sim Ber(p)$
Uniforme$(\{a, \dots, b\})$	(a, b)	$\frac{a+b}{2}$	$\frac{n^2-1}{12}$ $n = b-a+1$	—
Binomiale(n, p)	(n, p)	np	$np(1-p)$	$Bin(n_1+n_2, p)$ $X \sim Bin(n_1,p)$ $Y \sim Bin(n_2,p)$
Geometrica(p)	p	$1/p$	$(1-p)/p^2$	—
Poisson(λ)	λ	λ	λ	$Pois(\lambda_1+\lambda_2)$
Normale(μ, σ^2)	(μ, σ^2)	μ	σ^2	$\mathcal{N}(\mu + \mu, \sigma^2 + \sigma^2)$

Sia $(X_n)_{n \geq 0}$ una successione di variabili aleatorie indipendenti e identicamente distribuite con valore atteso $\mathbb{E}(X_i) = \mu = 9$ e $\mathrm{SD}(X_i) = \sigma = 3$. Poniamo $\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$.

(a) Approssimativamente, quanto ci aspettiamo che valga $\bar{X}_n$ per n molto grande? Motivare la propria risposta con un risultato visto in classe.

(b) Stimare con la disuguaglianza di Chebyshev il valore di

$$\mathbb{P}[|\bar{X}_n - 9| \geq 30]$$

a) Per la legge dei grandi numeri $\bar{X}_n$ converge per n molto grande verso $\mathbb{E}(X_i) = 9$

$$\lim_{n \to \infty} \mathbb{P}(|\bar{X}_n - 9| \geq \varepsilon) = 0$$

b) $$\mathbb{P}(|\bar{X}_n - 9| \geq 30) = \mathbb{P}\left(|\bar{X}_n - 9| \geq \frac{30}{3} \cdot \sigma\right) \overset{C}{\leq} \frac{1}{\left(\frac{30}{3}\right)^2} = \frac{1}{100}$$

3. Una fabbrica produce delle lampadine speciali. In media, il 2.3% delle lampadine prodotte ha un difetto di produzione. Un controllo qualità prende 1000 lampadine, scelte a caso tra quelle prodotte (quindi assunte indipendenti), e le testa per vedere se sono difettose. Usando il teorema del limite centrale, calcolare la probabilita che nel campione ci siano più di 28 lampadine difettose. Fare questo calcolo con e senza correzione di continuità.

$$S_{1000} = \#\text{ lampadine difettose su } 1000 \sim \mathrm{Bin}(\underset{n}{1000}, \underset{p}{0{,}023})$$

$$\overset{CLT}{\Longrightarrow} \quad S_{1000} \approx \mathcal{N}(\mu_\#, \sigma_\#^2) \quad \text{con} \quad \begin{cases} \mu_\# = n \cdot p = 23 \\ \sigma_\# = \sqrt{np(1-p)} \simeq 4.74 \end{cases}$$

correzione di continuità

$$\Longrightarrow \quad \mathbb{P}\left(S_{1000} \geq 29 - \tfrac{1}{2}\right) = \mathbb{P}\left(\underbrace{\frac{S_{1000} - \mu_\#}{\sigma_\#}}_{\approx \mathcal{N}(0,1) \text{ (standardizzata)}} \geq \frac{29 - \frac{1}{2} - \mu_\#}{\sigma_\#}\right)$$

$$= 1 - \mathbb{P}\left(\frac{S_{1000} - \mu_\#}{\sigma_\#} \leq \frac{29 - \frac{1}{2} - \mu_\#}{\sigma_\#}\right)$$

$$= 1 - \Phi\left(\frac{29 - 23 - \frac{1}{2}}{4.74}\right)$$

$$= \begin{cases} 1 - \Phi(1.26) = 0.896 \\ 1 - \Phi(1.16) = 0.877 \end{cases}$$