Probabilità e Statistica – Corso di Laurea in Informatica
A.A. 2020/2021

ESERCITAZIONE 10

E10.1. Io prendo mediamente due raffreddori all'anno. Il tempo tra la fine di un raffreddore e l'inizio del successivo è una variabile aleatoria gaussiana con media 160 giorni e deviazione standard di 40 giorni. Qual è la probabilità che (a) rimanga 200 giorni o più senza raffreddore? (b) Mi prenda un raffreddore entro 80 giorni dalla guarigione del precedente?

Soluzione. Sia X il tempo che intercorre tra due raffreddori successivi. Abbiamo $X \sim \mathrm{N}(160, 40^2)$. Poiché

$$Z = \frac{X - 160}{40} \sim \mathrm{N}(0,1)$$

calcoliamo

(a) $P(X \geq 200) = P\left(Z \geq \dfrac{200-160}{40}\right) = P(Z \geq 1) = 1 - \Phi(1) \simeq 1 - 0.8413 = 0.1587;$

(b) $P(X \leq 80) = P\left(Z \leq \dfrac{80-160}{40}\right) = P(Z \leq -2) = \Phi(-2) = 1 - \Phi(2) \simeq 1 - 0.9772 = 0.0228.$

E10.2. L'altezza media (in centimetri) di un bambino di 7 mesi è una variabile aleatoria normale di parametri $\mu = 71$ e $\sigma^2 = 6.25$.
(a) Qual è la percentuale di bambini di 7 mesi che superano i 74 centimetri di altezza? (b) Se si selezionano a caso tre bambini di 7 mesi, qual è la probabilità che tutti e tre superino i 74 centimetri di altezza?

Soluzione. Sia X l'altezza (in centimetri) di un bambino di 7 mesi. Abbiamo $X \sim \mathrm{N}(71, 6.25)$.

(a) Poiché

$$Z = \frac{X - 71}{\sqrt{6.25}} \sim \mathrm{N}(0,1)$$

si ha

$$P(X \geq 74) = P\left(Z \geq \frac{74-71}{\sqrt{6.25}}\right) = P(Z \geq 1.2) = 1 - \Phi(1.2) \simeq 1 - 0.8849 = 0.1151$$

e quindi circa il 11.5% dei bambini di 7 mesi supera i 74 centimetri di altezza.

(b) Trattandosi di tre eventi indipendenti, la probabilità richiesta è data da

$$[P(X \geq 74)]^3 = (0.1151)^3 \simeq 0.0015.$$

E10.3. Il quoziente d'intelligenza (QI) nella popolazione ha distribuzione normale con media 100 e deviazione standard 15. Quale valore minimo di QI bisogna avere per appartenere al 5% della popolazione con maggiore QI?

Soluzione. Sia $X_{💡}$ il valore del QI. Abbiamo $X_{💡} \sim \mathrm{N}(100, 15^2)$ e cerchiamo $x \in \mathbb{R}$ tale che $P\left(X_{💡} > x\right) = 0.05$. Poiché

$$Z = \frac{X_{💡} - 100}{15} \sim \mathrm{N}(0,1)$$

si ha

$$P\left(X_{💡} > x\right) = P\left(Z > \frac{x-100}{15}\right) = 1 - \Phi\left(\frac{x-100}{15}\right)$$

e quindi deve essere

$$1-\Phi\left(\frac{x-100}{15}\right) = 0.05 \quad \Leftrightarrow \quad \Phi\left(\frac{x-100}{15}\right) = 0.95 \quad \Leftrightarrow \quad \frac{x-100}{15} = 1.645 \quad \Leftrightarrow \quad x = 124.67\,.$$

E10.4. Siano $X, Y \sim \mathrm{N}(0, \sigma^2)$ indipendenti. Mostrare che $W := X - Y$ ha distribuzione $\mathrm{N}(0, 2\sigma^2)$. Inoltre calcolare $P(X \geq Y)$ e $P(X \geq Y + 1)$.

Soluzione.

- La variabile aleatoria W è una trasformazione lineare di variabili aleatorie gaussiane *indipendenti*, pertanto è anche lei una variabile aleatoria gaussiana. Determiniamone i parametri. Poiché si ha

$$E(W) \overset{\text{(linearità)}}{=} E(X-Y) = E(X) - E(Y) = 0 - 0 = 0$$

  e

$$Var(W) \overset{\text{(additività)}}{=} Var(X) + (-1)^2 Var(Y) = \sigma^2 + \sigma^2 = 2\sigma^2,$$

  otteniamo $W \sim \mathrm{N}(0, 2\sigma^2)$.

- Abbiamo $P(X \geq Y) = P(X - Y \geq 0) = P(W \geq 0)$. Poiché W è una variabile aleatoria gaussiana con valor medio 0, come conseguenza della simmetria della distribuzione, otteniamo $P(W \geq 0) = \frac{1}{2}$.

- Abbiamo $P(X \geq Y + 1) = P(X - Y \geq 1) = P(W \geq 1)$. Poiché

$$Z = \frac{W}{\sqrt{2\sigma^2}} \sim \mathrm{N}(0,1)$$

  si trova

$$P(W \geq 1) = P\left(Z \geq \frac{1}{\sqrt{2\sigma^2}}\right) = 1 - \Phi\left(\frac{1}{\sqrt{2\sigma^2}}\right).$$

E10.5 (difficile; video). Il peso corporeo nella popolazione maschile adulta ha distribuzione normale con media 72 kg e deviazione standard 13.2 kg, mentre in quella femminile ha distribuzione normale con media 64 kg e deviazione standard 10.7 kg. In un ascensore è esposta la portata dichiarata di 5 persone e 400 kg.
(a) Se nell'ascensore salgono 5 maschi, qual è la probabilità che il loro peso complessivo superi la portata dell'ascensore? (b) E se salgono 4 maschi e una femmina?

Soluzione.

(a) Siano $X_1^{♂}, \ldots, X_5^{♂}$ i pesi dei 5 maschi che salgono sull'ascensore. Abbiamo $X_1^{♂}, \ldots, X_5^{♂} \sim \mathrm{N}(72, 13.2^2)$ e sono tra loro *indipendenti.* Il peso complessivo è dato da

$$X = \sum_{i=1}^{5} X_i^{♂} \sim \mathrm{N}(5 \cdot 72, 5 \cdot 13.2^2) = \mathrm{N}(360, 871.2)\,.$$

Poiché

$$Z = \frac{X - 360}{\sqrt{871.2}} \sim \mathrm{N}(0,1)$$

otteniamo

$$P(X > 400) = P\left(Z > \frac{400-360}{\sqrt{871.2}}\right) = P(Z > 1.36) = 1 - \Phi(1.36) \simeq 1 - 0.9131 = 0.0869\,.$$

(b) Siano $X_1^{\male}, \ldots, X_4^{\male}$ i pesi dei 4 maschi e $X^{\female}$ il peso della femmina che salgono sull'ascensore. Abbiamo $X_1^{\male}, \ldots, X_4^{\male} \sim \mathrm{N}(72, 13.2^2)$, $X^{\female} \sim \mathrm{N}(64, 10.7^2)$ e sono tra loro *indipendenti*. Il peso complessivo è dato da

$$Y = \sum_{i=1}^{4} X_i^{\male} + X^{\female} \sim \mathrm{N}(4 \cdot 72 + 64, 4 \cdot 13.2^2 + 10.7^2) = \mathrm{N}(352, 811.45)\,.$$

Poiché

$$Z = \frac{Y - 352}{\sqrt{811.45}} \sim \mathrm{N}(0,1)$$

otteniamo

$$P(Y > 400) = P\left(Z > \frac{400 - 352}{\sqrt{811.45}}\right) = P(Z > 1.69) = 1 - \Phi(1.69) \simeq 1 - 0.9545 = 0.0455\,.$$

E10.6. Siano $X_1, X_2, \ldots, X_{625}$ variabili aleatorie indipendenti ed identicamente distribuite, con funzione di densità

$$f(x) = \begin{cases} 3(1-x)^2 & \text{se } 0 \le x \le 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Usare il teorema del limite centrale per approssimare $P(X_1 + X_2 + \cdots + X_{625} < 170)$.

Soluzione. Determiniamo valor medio e varianza delle X_i. Calcoliamo

$$E(X_i) = \int_{-\infty}^{+\infty} x f(x)\, dx = 3 \int_0^1 x(1-x)^2\, dx = 3\left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4}\right]_0^1 = \frac{1}{4}$$

e

$$Var(X_i) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx - \left(\frac{1}{4}\right)^2 = 3\int_0^1 x^2(1-x)^2\, dx - \frac{1}{16} = 3\left[\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5}\right]_0^1 - \frac{1}{16} = \frac{3}{80}.$$

Pertanto, usando l'approssimazione normale, otteniamo

Nota: 0,272 si ottiene facendo 170/625 (questo si fa per mantenere la dis.)

$$P(X_1 + \cdots + X_{625} < 170) = P\left(\overline{X}_{625} < 0.272\right) = P\left(25 \cdot \frac{\overline{X}_{625} - 0.25}{\sqrt{0.0375}} < 25 \cdot \frac{0.272 - 0.25}{\sqrt{0.0375}}\right)$$
$$\approx P(Z < 2.84) = \Phi(2.84) = 0.9977,$$

Qui sopra tutta la parte a sx sarà espressa come Z, mentre tutta la parte di dx sarà espressa come 2.84

dove $Z \sim \mathrm{N}(0,1)$.

E10.7. Un generatore di numeri casuali genera numeri compresi tra 0 e 1, secondo una distribuzione di media 0.5 e deviazione standard 0.289. Se vengono generati 100 numeri, qual è la probabilità (approssimata) che la loro media empirica sia compresa tra 0.47 e 0.53?

Soluzione. Sia X_i l'i-esimo numero casuale generato. Le variabili $X_1, \ldots, X_{100}$ sono indipendenti e identicamente distribuite con $E(X_i) = 0.5$ e $Var(X_i) = (0.289)^2$ per ogni $i = 1, \ldots, 100$. Usando l'approssimazione normale, otteniamo

$$P(0.47 \le \overline{X}_{100} \le 0.53) = P\left(\frac{0.47 - 0.5}{0.0289} \le \frac{\overline{X}_{100} - 0.5}{0.0289} \le \frac{0.53 - 0.5}{0.0289}\right)$$
$$\approx P(-1.04 \le Z \le 1.04) = 2\Phi(1.04) - 1 = 2 \cdot (0.8508) - 1 = 0.7016,$$

dove $Z \sim \mathrm{N}(0,1)$.

E10.8. Nella battitura di un testo ogni carattere viene sbagliato con probabilità 0.005, indipendentemente dagli altri. Un articolo contiene 2500 caratteri. Usando l'approssimazione normale, calcolare la probabilità che nell'articolo ci siano oltre 15 errori.

Soluzione. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 2500$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo carattere viene sbagliato} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, \dots, X_{2500}$ sono indipendenti e identicamente distribuite con distribuzione Be(0.005). In particolare, si ha $E(X_i) = 0.005$ e $Var(X_i) = (0.005)(0.995) = 0.004975$, usando l'approssimazione normale, otteniamo

Nota: rad(2500) = 50, cioè il numero scritto a sinistra del prodotto normale

$$P(X_1 + \dots + X_{2500} > 15) = P(\overline{X}_{2500} > 0.006) = P\left(50 \cdot \frac{\overline{X}_{2500} - 0.005}{0.07} > 50 \cdot \frac{0.006 - 0.005}{0.07}\right)$$
$$\approx P(Z > 0.71) = 1 - \Phi(0.71) = 1 - 0.7611 = 0.2389,$$

dove $Z \sim \mathrm{N}(0,1)$.

E10.9 (difficile; video). La probabilità di trovare una perla in un'ostrica non coltivata è pari a circa 0.001. Quante ostriche si dovranno pescare per poter disporre, con probabilità uguale a 0.3, di almeno 10 perle?

Soluzione. Costruiamo uno schema di Bernoulli. Per $n \geq 1$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se nell'}i\text{-esima ostrica c'è una perla} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, X_2, \dots$ sono indipendenti e identicamente distribuite con distribuzione Be(0.001). In particolare, si ha $E(X_i) = 0.001$ e $Var(X_i) = (0.001)(0.999) \approx 0.001$. L'incognita è il numero n di perle che vanno pescate. Per ricavare n dobbiamo risolvere $P(X_1 + \dots + X_n \geq 10) = 0.3$. Abbiamo

$$P(X_1 + \dots + X_n \geq 10) = P\left(\overline{X}_n \geq \tfrac{10}{n}\right) = P\left(\sqrt{n} \cdot \frac{\overline{X}_n - 0.001}{0.032} \geq \sqrt{n} \cdot \frac{\frac{10}{n} - 0.001}{0.032}\right)$$
$$\approx P\left(Z \geq \frac{10 - 0.001n}{0.032\sqrt{n}}\right) = 1 - \Phi\left(\frac{10 - 0.001n}{0.032\sqrt{n}}\right),$$

dove $Z \sim \mathrm{N}(0,1)$. Quindi, otteniamo l'equazione

$$\begin{aligned} 1 - \Phi\left(\frac{10 - 0.001n}{0.032\sqrt{n}}\right) = 0.3 \quad &\Leftrightarrow \quad \Phi\left(\frac{10 - 0.001n}{0.032\sqrt{n}}\right) = 0.7 \\ &\Leftrightarrow \quad \frac{10 - 0.001n}{0.032\sqrt{n}} = 0.525 \\ &\Leftrightarrow \quad 0.001n + 0.0168\sqrt{n} - 10 = 0 \\ &\Leftrightarrow \quad \sqrt{n} \approx 91.95 \text{ o } \sqrt{n} \approx -108.75 \text{ (non accettabile)}, \end{aligned}$$

da cui segue $n = 8456$.

E10.10 (tratto da appello). Mia nonna era una fanatica di concorsi televisivi (non sto scherzando). Tra i suoi preferiti, c'erano i giochi a premi della trasmissione *I fatti vostri*. Chiamava e richiamava ad oltranza il numero `894.433` nella speranza di prendere la linea, lasciare i suoi dati e poi venir sorteggiata per giocare.

(a) Se ad ogni telefonata, indipendentemente dalle altre, la probabilità di prendere la linea è pari a 0.025, qual è la probabilità che mia nonna ci riuscisse entro la sesta chiamata? (b) Supponiamo ci siano 150 telespettatori che telefonano nel tentativo di partecipare al gioco a premi. Qual è la probabilità (approssimata) che più di 30 tra loro riescano a prendere la linea entro la sesta chiamata?

Soluzione.

(a) Costruiamo uno schema di Bernoulli. Per $i \geq 1$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se la nonna prende la linea all'}i\text{-esima chiamata} \\ 0 & \text{altrimenti.} \end{cases}$$

Le variabili $X_1, X_2, \ldots$ sono indipendenti ed identicamente distribuite con distribuzione $\mathrm{Be}(0.025)$. Siamo interessati alla prima chiamata con cui la nonna riesce a prendere la linea (istante di primo successo), cioè alla variabile aleatoria $T = \inf\{i \geq 1 : X_i = 1\} \sim \mathrm{Ge}(0.025)$. Pertanto, otteniamo

$$P(T \leq 6) = 1 - P(T > 6) = 1 - (1 - 0.025)^6 \approx 0.14,$$

dove abbiamo usato la formula per la probabilità di lunga attesa.

(b) Costruiamo un altro schema di Bernoulli. Per $i = 1, \ldots, 150$, definiamo le variabili aleatorie

$$Y_i = \begin{cases} 1 & \text{se l'}i\text{-esimo telespettatore prende la linea entro la sesta chiamata} \\ 0 & \text{altrimenti.} \end{cases}$$

Le variabili $Y_1, \ldots, Y_{150}$ sono indipendenti ed identicamente distribuite con distribuzione $\mathrm{Be}(0.14)$. In particolare, si ha $E(Y_i) = 0.14$ e $Var(Y_i) = (0.14)(0.86) = 0.1204$. Calcoliamo

$$P(Y_1 + \cdots + Y_{150} > 30) = P(\overline{Y}_{150} > 0.2) = P\left(\sqrt{150} \cdot \frac{\overline{Y}_{150} - 0.14}{\sqrt{0.1204}} > \sqrt{150} \cdot \frac{0.2 - 0.14}{\sqrt{0.1204}}\right)$$

$$\approx P(Z > 2.12) = 1 - \Phi(2.12) = 1 - 0.9830 = 0.017,$$

dove $Z \sim \mathrm{N}(0, 1)$.

E10.11 (⚑ tratto da appello; 🎞 video). All'ufficio postale `Padova 11` è aperto un solo sportello e il tempo (in minuti) che un utente ci passa per essere servito è una variabile aleatoria X con distribuzione esponenziale di media 10. Calcolare la probabilità che (a) un utente stia allo sportello più di 15 minuti; (b) un utente stia allo sportello più di 15 minuti, sapendo che ci ha già passato 10 minuti.

Mi reco all'agenzia `Padova 11` per ritirare una raccomandata. Al momento del mio arrivo sono la decima persona in fila e la prima sta andando ora allo sportello. Ogni utente, indipendentemente dagli altri, passerà allo sportello un tempo (in minuti) distribuito come una variabile aleatoria esponenziale di media 10. Usando l'approssimazione normale, calcolare la probabilità che (c) io riesca a ritirare la raccomandata entro un'ora.

Soluzione.

(a) Dal momento che $X \sim \mathrm{Exp}\left(\frac{1}{10}\right)$, otteniamo $P(X > 15) = 1 - F_X(15) = e^{-3/2}$.

(b) Per la proprietà di assenza di memoria, si ha

$$P(X > 15 \,|\, X > 10) = P(X > 5) = 1 - F_X(5) = e^{-1/2}.$$

(c) Sia A l'evento "riesco a ritirare la raccomandata entro un'ora". Inoltre, per $i \in \{1, \dots, 10\}$, sia X_i la variabile aleatoria che quantifica il tempo (in minuti) passato dall'i-esimo utente allo sportello della posta. Le variabili aleatorie $X_1, \dots, X_{10}$ sono indipendenti ed identicamente distribuite con distribuzione $\mathrm{Exp}\left(\frac{1}{10}\right)$. In particolare, si ha $E(X_i) = 10$ e $Var(X) = 100$. Calcoliamo

$$P(A) = P(X_1 + \cdots + X_{10} \leq 60) = P(\overline{X}_{10} \leq 6) = P\left(\sqrt{10} \cdot \frac{\overline{X}_{10} - 10}{10} \leq \sqrt{10} \cdot \frac{6 - 10}{10}\right)$$

$$\approx P(Z \leq -1.27) = \Phi(-1.27) = 1 - \Phi(1.27) = 1 - 0.8980 = 0.102,$$

dove $Z \sim \mathrm{N}(0, 1)$.