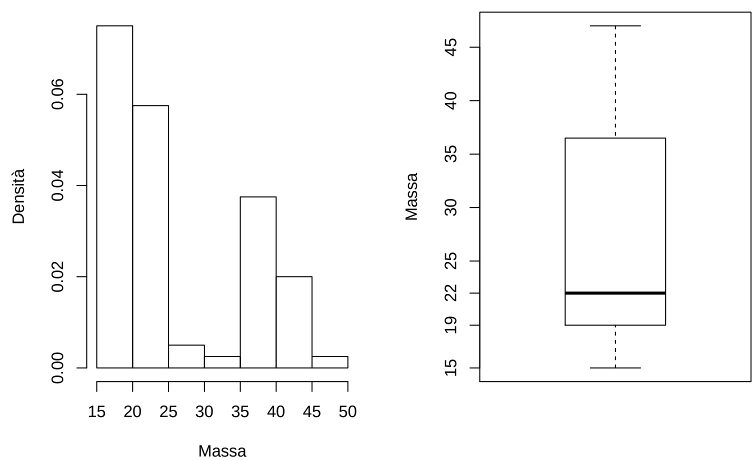
Politecnico di Milano - Scuola di Ingegneria Industriale e dell'Informazione

I Prova in Itinere di Statistica per Ingegneria Fisica
26 aprile 2017

Cognome, Nome, Matricola:

Problema 1. Nelle miniere di Moria si trovano ricchissimi giacimenti di oro, argento e pietre preziose. Due squadre di minatori, al comando di Balin, stanno esplorando i siti dei vecchi scavi per scoprire quali fossero i più produttivi. I minatori raccolgono e pesano 80 diamanti, provenienti da due scavi differenti. La distribuzione delle masse trovate (in grammi) è rappresentata dai grafici seguenti.

Distribuzione della massa dei diamanti



1. Commentare la distribuzione della massa dei diamanti raccolti.
2. Quanto vale la mediana dei dati?
3. Quanti diamanti hanno massa compresa tra 19 e 22 grammi?

(SEGUE)

I minatori si accorgono che i 25 diamanti più grandi provengono tutti da un unico sito, e vogliono approfondire lo studio delle relative masse (in grammi):

34, 36, 36, 36, 36, 37, 37, 38, 38, 39, 39, 39, 40, 40, 40, 40, 41, 41, 41, 41, 42, 43, 43, 44, 47.

4. Calcolare la mediana dei dati.
5. Determinare l'intervallo che contiene il 50% "centrale" dei dati (quelli appartenenti alla scatola di un boxplot) e calcolarne la lunghezza.
6. Ci sono valori erratici?
7. Completare la seguente tabella di distribuzione di frequenza dei dati

Classe	F. ass.	F. rel.	F. rel. cumulate	Densità
$(33, 36]$				
$(36, 39]$				
$(39, 42]$				
$(42, 45]$				
$(45, 48]$				

8. Disegnare l'istogramma relativo alla tabella calcolata e commentare la distribuzione dei dati.

Risultati.

1. Dall'esame dell'istogramma la distribuzione dei dati appare chiaramente bimodale, il che può essere conseguenza della provenienza dei diamanti da due distinti siti produttivi. Immaginiamo quindi che la provenienza sia una variabile latente che spiega la stratificazione in gruppi dei dati. Il boxplot conferma questa analisi e mostra una grande asimmetria.
   Dall'esame combinato di istogramma e boxplot, sembra che la numerosità dei dati nei due potenziali gruppi evidenziati dall'istogramma sia differente.

2. Dal boxplot si legge che la mediana vale 22 grammi.

3. Dall'esame del boxplot, si vede che circa il 25% dei diamanti (ossia circa 20) ha massa compresa tra 19 e 22 grammi.

4. La mediana campionaria vale: $\widehat{Q}_2 = \widehat{q}_{0.5} = 40$ grammi.

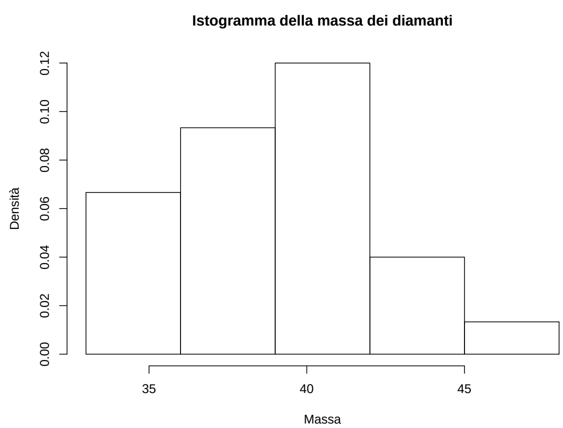
5. Il primo quartile vale $\widehat{Q}_1 = \widehat{q}_{0.25} = 37$ grammi, mentre il terzo quartile vale $\widehat{Q}_3 = \widehat{q}_{0.75} = 41$ grammi. L'intervallo desiderato è $IQR = [37, 41]$ grammi e ha lunghezza 4 grammi.

6. Sono erratici i dati inferiori a $37 - 6 = 31$ e quelli superiori a $41 + 6 = 47$. Non ci sono quindi valori erratici.

7. Tabella di distribuzione di frequenza dei dati:

Classe	F. ass.	F. rel.	F. rel. cumulate	Densità
$(33, 36]$	5	0.2	0.2	0.067
$(36, 39]$	7	0.28	0.48	0.093
$(39, 42]$	9	0.36	0.84	0.12
$(42, 45]$	3	0.12	0.96	0.04
$(45, 48]$	1	0.04	1	0.013

8. Istogramma:



Si nota come la bimodalità sia scomparsa, mentre questa scelta delle classi evidenzia una leggera asimmetria dei dati.

Problema 2. Nel prossimo anno scolastico ad Hogwarts Barty Crouch deve impersonare il professor Moody fra l'ora di colazione, le 7.00, e l'ora di cena, le 18.00. Sta quindi valutando di usare la pozione Polisucco con cui potrà assumere le sembianze del professore mentre si trova in pubblico. Sui libri di Pozioni, Barty Crouch scopre che la pozione può essere assunta una sola volta al giorno e che la durata T dei suoi effetti è aleatoria, con distribuzione normale di media 12.5 ore e deviazione standard 1 ora. Una volta svanito l'effetto, Barty Crouch si ritrasformerà subito in se stesso. Assumendo quindi la pozione alle ore 7.00, non si può escludere che venga scoperto prima delle 18.00.

1. Con quale probabilità Barty Crouch può venire scoperto tra colazione e cena?

Sui libri di Pozioni, Barty Crouch scopre anche che le durate di diversi sorsi di Polisucco sono indipendenti.

2. Con quale probabilità Barty Crouch non verrà scoperto mai durante una settimana (7 giorni)?
3. Con quel probabilità verrà scoperto al massimo 10 volte durante l'anno scolastico (200 giorni)?

In alternativa alla pozione Polisucco, Barty Crouch potrebbe bere la pozione Polisucco Light. La durata W del suo effetto è mediamente inferiore (è normale di media 7.5 ore e deviazione standard 1 ora), ma in compenso questa pozione può essere assunta due volte al giorno: a colazione, alle 7.00, e a pranzo, alle 12.00. Anche le durate di diversi sorsi di Polisucco Light sono indipendenti e, se a pranzo Barty Crouch beve il secondo sorso di Polisucco Light prima di essersi ritrasformato in se stesso, allora l'effetto del primo sorso si annulla e inizia l'effetto del secondo sorso.

4. Con quale probabilità Barty Crouch può venire scoperto tra colazione e cena, bevendo Polisucco Light a colazione e pranzo?

Risultati.

1. Sia T la durata della pozione Polisucco, per cui $T \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu = 12.5$ e $\sigma = 1$. Quindi

$$\mathbb{P}(T < 11) = \Phi\left(\frac{11 - 12.5}{1}\right) = 0.0668,$$

2. Sia X il numero di giorni su 7 in cui Barty Crouch sarà scoperto. Essendo indipendenti le durate della pozione nei diversi giorni, abbiamo $X \sim B(7, p)$ con $p = 0.0668$. Quindi

$$\mathbb{P}\left(X = 0\right) = \binom{7}{0} p^0 (1-p)^7 = 0.6163.$$

3. Sia Y il numero di giorni su 200 in cui Barty Crouch sarà scoperto. Essendo indipendenti le durate della pozione nei diversi giorni, abbiamo $Y \sim B(200, p)$ con $p = 0.0668$. Quindi

$$\mathbb{P}(Y \leq 10) = \sum_{k=0}^{10} \binom{200}{k} p^k (1-p)^{200-k} = 0.2128$$

Essendo $\mathbb{E}[Y] = n \cdot p = 200p = 13.36 > 5$ e $n - \mathbb{E}[Y] = n \cdot (1-p) = 200(1-p) = 186.64 > 5$, possiamo anche usare l'approssimazione normale $Y \sim B(200, p) \simeq N(13.36, 12.467552)$. Quindi

$$\mathbb{P}(Y \leq 10) = \mathbb{P}\left(\frac{Y - 13.36}{\sqrt{12.467552}} \leq \frac{10.5 - 13.36}{\sqrt{12.467552}}\right) \simeq \Phi(-0.81) = 0.2090$$

4. Sia $W_1 \sim \mathcal{N}(\mu, \sigma^2)$ la durata del Polisucco Light assunto a colazione, e sia $W_2 \sim \mathcal{N}(\mu, \sigma^2)$ la durata del Polisucco Light assunto a pranzo, con $\mu = 7.5$ e $\sigma = 1$. Allora

$$P(W_1 < 5) = \Phi\left(\frac{5 - 7.5}{1}\right) = 0.0062,$$
$$P(W_2 < 6) = \Phi\left(\frac{6 - 7.5}{1}\right) = 0.0668.$$

Sia X la v.a. che indica se Barty Crouch viene scoperto durante l'intero arco della giornata (cioè, dalle 07.00 alle 18.00). Allora è $X \sim B(p)$, con p da determinare. In particolare $X = 1$ se Barty Crouch viene scoperto almeno una volta fra mattina e pomeriggio, mentre $X = 0$ se Barty Crouch **non** viene scoperto di mattina e **non** viene scoperto di pomeriggio. Allora, sapendo che W_1 e W_2 sono indipendenti, troviamo:

$$\begin{aligned} p = \mathbb{P}\Big(X = 1\Big) = 1 - \mathbb{P}\Big(X = 0\Big) = 1 - \mathbb{P}\Big(W_1 \geq 5, W_2 \geq 6\Big) &= 1 - \mathbb{P}\Big(W_1 \geq 5\Big)\mathbb{P}\Big(W_2 \geq 6\Big) \\ &= 1 - (1 - 0.006)(1 - 0.067) = 0.0726. \end{aligned}$$

Problema 3. L'inverno è arrivato... e Sam sa che la durata T di un inverno è casuale con distribuzione esponenziale. Inoltre Sam scopre nella Biblioteca della Cittadella che 1 inverno su 2 dura più di 5 anni.

1. Quanto dura mediamente un inverno?
2. Con quale probabilità l'inverno appena iniziato durerà più di 9 anni?
3. Daenerys all'inizio dell'inverno intraprende una traversata del Mare dell'Estate e del Mare Stretto che durerà 1 mese. Con quale probabilità, compiuta la traversata, l'inverno sarà già finito?
4. Se invece Sam si sbagliasse e la durata T di un inverno avesse distribuzione uniforme, fra 0 ed un qualche limite superiore, con quale probabilità l'inverno appena iniziato potrebbe durare più di 9 anni? Nessuno mette invece in dubbio quanto scoperto da Sam nella Biblioteca della Cittadella.

Risultati. Misuriamo il tempo in anni.

1. $T \sim \mathcal{E}(\lambda)$ con $\mathbb{P}(T > 5) = 0.5$. Si ricava
$$\mathbb{P}(T > 5) = \mathrm{e}^{-5\lambda} = \frac{1}{2} \qquad \Rightarrow \qquad \lambda = \frac{1}{5}\log 2 = 0.1386$$
per cui
$$\mathbb{E}[T] = \frac{1}{\lambda} = \frac{5}{\log 2} = 7.21$$
ovvero 7 anni, 2 mesi e 17 giorni.

2. $\mathbb{P}(T > 9) = \mathrm{e}^{-9\lambda} = \mathrm{e}^{-\frac{9\log 2}{5}} = 0.2872.$

3. $\mathbb{P}\left(T < \frac{1}{12}\right) = 1 - \mathrm{e}^{-\lambda/12} = 1 - \mathrm{e}^{-\frac{\log 2}{60}} = 0.0115.$

4. Supponiamo ora $T \sim U(0, b)$, ma sempre con $\mathbb{P}(T > 5) = 0.5$. Si ricava
$$\frac{b-5}{b} = 0.5 \qquad \Rightarrow \qquad b = 10$$
per cui
$$\mathbb{P}(T > 9) = \frac{1}{10} = 0.1.$$