Calcolo delle Probabilità e Statistica
Soluzioni degli esami

Magoga Francesco

12 febbraio 2019

Versione 1.6

Indice

Introduzione

Le soluzioni contenute in questo testo si basano sugli esami dell'A.A. 2016/2017, sull'A.A. 2017/2018 e sull'inizio dell'A.A. 2018/2019 del corso di Calcolo delle Probabilità e Statistica tenuto dal prof. Luigi Pace all'università degli studi di Udine.

Le soluzioni dovrebbero essere tutte corrette. In caso trovaste degli errori potete contattarmi sulla mail spes (magoga.francesco@spes.uniud.it), così che io possa correggerli e rilasciare una nuova versione della dispensa (si, la versione serve solo per capire quale dispensa è scritta più recentemente, i contenuti però non cambiano, a meno di correzioni di errori).

Le soluzioni sono spiegate esaustivamente, ma per capirle è necessaria una conoscenza di base della materia. In certi casi ho preferito scrivere i procedimenti di come si ricavano alcune formule da utilizzare, è poi a discrezione vostra riportarli nell'esame o meno.

1 Urne e palline

1.1 Testo

Un'urna contiene N_1 palline nere e B_1 bianche. Una seconda urna contiene N_2 palline nere e B_2 bianche. Una terza urna contiene N_3 palline nere e B_3 bianche. (Una quarta urna contiene N_4 palline nere e B_4 bianche.) Uno sperimentatore sceglie a caso un'urna fra le n con equiprobabilità, poi estrae a caso, con reinserimento, n_E palline dall'urna scelta. Si determini la probabilità che l'urna scelta sia stata quella con N_j nere, se le palline estratte risultano, senza tener conto dell'ordine di estrazione, una N_E e B_E bianche.

1.2 Soluzione

Abbiamo n urne A_i con $i = 1, \ldots, n$ tali che l'urna A_i contiene N_i palline nere e B_i palline bianche per un totale di $n_i = N_i + B_i$ palline. Sappiamo che è stata scelta un'urna A_j con $1 \leq j \leq n$ e l'evento E ci dice che sono state estratte N_E palline nere e B_E palline bianche, per un totale di $n_E = N_E + B_E$.

Bisogna scegliere un colore di riferimento, che noi chiameremo con C, avremo che quindi $C_i = N_i$ se si è scelto il nero, $C_i = B_i$ se si è scelto il bianco. Sia $\overline{C_i}$ l'altro colore. Sia anche C_E il numero delle palline del colore scelto presenti in E. Sia $P_i(c) = \frac{c_i}{n_i}$ con $c = \{C_i, \overline{C_i}\}$ la probabilità di estrarre il colore c dall'urna i.

Per il compito: *Bisogna scrivere gli eventi! Quindi A_i = "viene scelta l'urna A_i" e E = "vengono estratte con reinserimento N_E palline nere e B_E palline bianche".*

Per il compito: *Bisogna dire che gli eventi A_i formano una partizione.*

Le verosimiglianze che indicano la probabilità che si sia verificato l'evento E avendo scelto l'urna A_i sono:

$$
\begin{aligned}
P(E \mid A_i) &= \binom{n_E}{C_E} P_i(C_i)^{C_E} P_i(\overline{C_i})^{(n_E - C_E)} \\
&= \binom{n_E}{C_i} \left(\frac{C_i}{n_i}\right)^{C_E} \left(\frac{\overline{C_i}}{n_i}\right)^{(n_E - C_E)}
\end{aligned}
$$

Per il compito: *Indicare che le probabilità sono binomiali poiché l'estrazione avviene con reinserimento delle palline nell'urna.*

Il coefficiente binomiale esprime la probabilità di estrarre C_E dello stesso colore da un numero di palline n_E, la prima potenza indica la probabilità dei casi di successo e la seconda quella dei casi di fallimento.

Adesso dobbiamo applicare Bayes per trovare la probabilità che sia verificato l'evento avendo scelto l'urna A_j.

$$P(A_j \mid E) = \frac{P(A_j)P(E \mid A_j)}{\sum_{i=1}^{n} P(A_i)P(E \mid A_i)} = \frac{\frac{1}{n}P(E \mid A_j)}{\frac{1}{n}\sum_{i=1}^{n} P(E \mid A_i)} = \frac{P(E \mid A_j)}{\sum_{i=1}^{n} P(E \mid A_i)}$$

Per il compito: *Indicare esplicitamente che si utilizza il teorema di Bayes.*

$P(A_i) = \frac{1}{n}$ dato che la scelta dell'urna è equiprobabile, poiché fatta a caso.

Per il compito: *Indicare esplicitamente che la probabilità* $P(A_i) = \frac{1}{n}$ *poiché l'estrazione è equiprobabile.*

Nota: nel secondo passaggio si è sostituito $P(A_i)$ e si è portata fuori la frazione dalla sommatoria.

2 V.c. continue e funzione di ripartizione

2.1 Testo

Sia X una variabile casuale con supporto $S_X = [a, b]$ e funzione di densità di probabilità di forma $p_X(x) = cf(x)$ per $x \in S_X$ e 0 altrove ($\exp(z) = e^z$). Si completi la definizione della funzione di densità di X, determinando il valore della costante di normalizzazione c. Si calcoli la funzione di ripartizione di X, esplicitandola in tutti i suoi tratti. Si ottengano valore atteso/mediana/moda di X. Sia infine $T = g(X)$. Si ottenga il supporto di T e si calcoli $P(h(T))$/si calcoli la varianza di T.

2.2 Soluzione

Abbiamo una v.c. continua definita su un intervallo $[a, b]$ con legge di probabilità contenente una costante c. Per calcolare tale costante basta integrare secondo la formula seguente:

$$\int_a^b p_X(x)dx = 1$$
$$\int_a^b cf(x)dx = 1$$
$$c\int_a^b f(x)dx = 1$$
$$c = \frac{1}{\int_a^b f(x)dx}$$

Nota: in realtà bisognerebbe integrare in $]-\infty, +\infty[$, ma sappiamo che in $]-\infty, a[$ e in $]b, +\infty[$ la v.c. non è definita quindi la p_X sarà sempre 0.

Esplicitando la funzione in tutti i tratti abbiamo che:

$$p_X(x) = \begin{cases} 0 & \text{se } x \notin [a, b] \\ cf(x) & \text{se } x \in [a, b] \end{cases} = \begin{cases} 0 & \text{se } x < a \vee x > b \\ \frac{1}{\int_a^b f(x)dx} f(x) & \text{se } a \leq x \leq b \end{cases}$$

La funzione di ripartizione è definita come $F_X(x) = P(X \leq x)$. Dato che la v.c. è continua $P(X = x) = 0$ quindi $F_X(x) = P(X < x)$. La funzione di ripartizione è definita come:

$$F_X(x) = \int_{-\infty}^x p_X(t)dt = \int_a^x p_X(t)dt = c\int_a^x f(t)dt$$

Possiamo ovviamente sostituire il $-\infty$ con a, dato che la v.c. nell'intervallo $]-\infty, a[$ non è definita e quindi $P(X \in]-\infty, a[) = 0$.

Esplicitando anche i tratti non compresi nell'intervallo avremo che:

$$F_X(x) = \begin{cases} 0 & \text{se } x \in \,]-\infty, a[\\ c\int_a^x f(t)dt & \text{se } x \in [a,b] \\ 1 & \text{se } x \in \,]b, +\infty[\end{cases} = \begin{cases} 0 & \text{se } x < a \\ c\int_a^x f(t)dt & \text{se } a \leq x \leq b \\ 1 & \text{se } x > b \end{cases}$$

Il valore atteso è la classica media aritmetica:

$$E(X) = \int_{-\infty}^{+\infty} xp_X(x)dx = \int_a^b xp_X(x)dx = c\int_a^b xf(x)dx$$

La forma in cui è scritta è strana, ma prendendo il valore $x \in S_X$ con probabilità $p_X(x) = v/n$ avremo che il valore x verrà ripetuto v volte e la divisione per n (fattore che verrà raccolto) è esattamente il numero di occorrenze dei valori $\in S_X$ e quindi $n = |S_X|$.

Varianza è *una misura della variabilità dei valori assunti dalla variabile stessa; nello specifico, la misura di quanto essi si discostino quadraticamente rispettivamente dalla media aritmetica o dal valore atteso* $E(X)$ (fonte: Wikipedia). Si calcola facilmente tramite la formula

$$Var(X) = E(X^2) - E(X)^2$$

dove $E(X^2)$ è il valore atteso di X dove tutte le occorrenze di x sono elevate al quadrato e $E(X)^2$ è valore atteso elevato al quadrato di X. In formule sarebbe:

$$\begin{aligned} Var(X) &= c\int_a^b x^2 f(x^2)dx - \left(c\int_a^b xf(x)dx\right)^2 \\ &= c\left[\int_a^b x^2 f(x^2)dx - c\left(\int_a^b xf(x)dx\right)^2\right] \end{aligned}$$

La moda è l'elemento del supporto che ha più probabilità di capitare e che effettivamente capita più frequentemente (è un trend):

$$x_{mo} \in S_X \ : \ p_X(x_{mo}) \geq p_X(x) \ \forall x \in S_X$$

Per calcolarla conviene osservare se la funzione $p_X(x)$ è monotona crescente o decrescente: se è crescente la moda sarà $x_{mo} = b$ poiché per definizione di funzione monotona crescente $p_X(a) < \cdots < p_X(b)$; se è decrescente la moda sarà $x_{mo} = a$ poiché per definizione di funzione monotona decrescente $p_X(a) > \cdots > p_X(b)$.

La mediana è l'elemento del supporto che ha probabilità esattamente in mezzo (quindi che spacca a metà) all'intervallo delle probabilità

$$x_{0,5} \in S_X \ : \ P_X(X \leq x_{0,5}) = P_X(X \geq x_{0,5}) = 0,5 = \frac{1}{2}$$

Per calcolarla si usa il primo termine:

$$P_X(X \leq x_{0,5}) = F_X(x) = \frac{1}{2}$$

Basta quindi porre la funzione di ripartizione uguale a $0,5$ e esplicitare la variabile x per calcolare la mediana.

Per il compito: *dare la definizione di ciò che è richiesto (valore atteso, varianza, moda, mediana).*

Per ultima cosa abbiamo una v.c. trasformata T, che deriva da X tramite una funzione g (i.e. $T = g(X)$ e $g(x) = -x$ oppure $g(x) = x^2$ o ancora $g(x) = |x|, \dots$). Il supporto di T equivale a:

$$S_T = g(S_X) = \{t \in \mathbb{R} : t = g(x) \text{ per qualche } x \in S_X\}$$

ovvero applicare la funzione g ad ogni singolo elemento di S_X.

Nota: sarebbe $t \in \mathbb{R}$ con d il grado della v.c. X che in questo caso $d = 1$ poiché X è univariata.

Dal supporto di t S_T possiamo dedurre che la v.c. trasformata T di una v.c. continua X è anch'essa continua. Questo implica che la probabilità in un singolo punto è pari a 0 (i.e. $P(T = c) = 0$ con $c \in S_T$). Negli altri casi è facile calcolare la probabilità tramite la funzione di ripartizione ricordando che $F_T(t) = P(T \leq t) = P(T < t)$ e $P(T \geq t) = P(T > t) = 1 - F_T(t)$:

$$F_T(t) = \begin{cases} F_X(g^{-1}(t)) & \text{se } g \text{ è monotona crescente} \\ 1 - F_X(g^{-1}(t)) & \text{se } g \text{ è monotona decrescente} \end{cases}$$

dove in realtà se g è decrescente $F_T(t) = 1 - F_X(g^{-1}(t)) + P(X = g^{-1}(t))$, ma la formula si semplifica dato che $P(X = g^{-1}(t)) = p_X(g^{-1}(t)) = 0$.

Si ricorda che per trovare $g^{-1}(t)$ basta risolvere l'equazione $t = g(x)$ in base alla variabile x (quindi esplicitare la variabile x) per ottenere $t = g^{-1}(x)$.

In qualche occasione al posto della probabilità di t viene richiesta la varianza di T. Per calcolarla usiamo la proprietà di linearità del valore atteso: sia $T = aX + b$ allora

$$E(T) = E(aX + b) = aE(X) + b$$

Per il compito: *Indicare esplicitamente che si usa la proprietà di linearità del valore atteso.*

La formula per calcolare la varianza di T la otteniamo così:

$$
\begin{aligned}
Var(T) &= E(T^2) - E(T)^2 \\
&= E((aX+b)^2) - (E(aX+b))^2 \\
&= E(a^2X^2 + aXb + b^2) - (aE(X)+b)^2 \\
&= a^2E(X^2) + aE(X)b + b^2 - (a^2E(X)^2 + aE(X)b + b^2) \\
&= a^2E(X^2) + aE(X)b + b^2 - a^2E(X)^2 - aE(X)b - b^2 \\
&= a^2E(X^2) - a^2E(X)^2 \\
&= a^2(E(X^2) - E(X)^2) \\
&= a^2Var(X)
\end{aligned}
$$

Quindi per calcolare la varianza di T basta calcolare la varianza di X e moltiplicarla per il termine a^2.

3 Binomiali e v.c. bivariate

3.1 Testo

Sia (X, Y) una variabile casuale bivariata con componente marginale $X \sim Bi(n_X, p_X)$ (legge binomiale con indice $n = n_X$ e parametro $p = p_X$) e distribuzioni condizionate binomiali $Y|X = x \sim Bi(1 + x, 1/2)$, per $x \in S_X$. Si determinino il supporto congiunto di (X, Y), la funzione di probabilità congiunta di (X, Y), il supporto marginale di Y, la funzione di probabilità marginale di Y. Si dica, motivando, se (X, Y) ha componenti indipendenti. Si calcoli infine $P(f(X, Y))$.

3.2 Soluzione

Abbiamo le v.c. $X = Bi(n_X, p_X)$ e $Y = Bi(n_Y, p_Y)$ entrambe con legge binomiale e Y dipendente da X.

Il supporto di X è determinato da:

$$S_X = \{x : \binom{n_X}{x}(p_X)^x(1 - p_X)^{(n_X - x)} > 0\} = \{0, \dots, n_X\}$$

Se Y dipende da X (i.e. in n_X oppure in p_X compare x) allora il supporto di Y si determina fissando un valore di $x \in S_X$:

$$S_{Y|X=x} = \{y : \binom{n_{Y|x}}{y}(p_{Y|x})^y(1 - p_{Y|x})^{(n_{Y|x} - y)} > 0\} = \{0, \dots, n_{Y|x}\}$$

dove i termini $n_{Y|x}$ e $p_{Y|x}$ sono i termini n_Y e p_Y nelle cui espressioni si è sostituita la variabile x con il valore fissato.

Se invece Y non dipende da X (i.e. in n_X E in p_X NON compare x) allora $S_{Y|X=x}$ si può determinare semplicemente come S_X:

$$S_{Y|X=x} = \{y : \binom{n_Y}{y}(p_Y)^y(1 - p_Y)^{(n_Y - y)} > 0\} = \{0, \dots, n_Y\}$$

Il supporto congiunto è quindi l'insieme di coppie:

$$S_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x \in S_X, y \in S_{Y|X=x}\}$$

Nota: se le v.c. sono dipendenti e supponiamo che $S_X = \{0, 1\}$, $S_{Y|X=0} = \{0, 1\}$ e $S_{Y|X=1} = \{1, 2\}$ avremo che $S_{X,Y} = \{(0, 0), (0, 1), (1, 1), (1, 2)\}$. Credo che il prof. richieda la scrittura esplicita di tutte le coppie.

La funzione di probabilità congiunta va determinata assegnando ogni $x \in S_X$ e ogni $y \in S_{Y|X=x}$:

$$p_{X,Y}(x, y) = p_X(x)p_{Y|X=x}(y)$$

La funzione $p_{Y|X=x}$ sarà determinata dalla legge $Bi(n_{Y|X=x}, p_{Y|X=x})$ che si ottiene sostituendo la variabile x con un preciso valore in S_X nella legge $Bi(n_Y, p_Y)$.

Si ricorda che

$$p_X(x) = \binom{n_X}{x}(p_X)^x(1-p_X)^{(n_X-x)}$$

e

$$p_{Y|X=x}(y) = \binom{n_{Y|x}}{y}(p_{Y|x})^y(1-p_{Y|x})^{(n_{Y|x}-y)}$$

che sono le solite probabilità binomiali.

Il supporto marginale di Y è:

$$S_Y = \{y \in \mathbb{R} : (x,y) \in S_{X,Y} \text{ per qualche } x\}$$

Quindi è l'insieme di tutti i secondi elementi delle coppie nel supporto congiunto.

La legge marginale va calcolata $\forall y \in S_Y$ nel seguente modo:

$$p_Y(y) = \sum_{x \in S_x} p_{X,Y}(x,y)$$

La v.c. bivariata ha componenti indipendenti se $S_{X,Y} = S_X \times S_Y$ e per ogni $(x,y) \in S_{X,Y}$ vale che $p_{X,Y}(x,y) = p_X(x)p_Y(y)$.

L'ultimo punto si ottiene sommando $p_{X,Y}(x,y)$ con x e y che soddisfano la condizione (i.e. se $X+Y=2$ prendo tutti gli (x,y) tali che $x+y=2$)

4 Resistenze

4.1 Testo

Una apparecchiatura ha solo n componenti che si possono guastare. La vita operativa X_i $(i = 1, ..., n)$ di ciascuna di esse ha distribuzione esponenziale *con valore atteso pari a* $E(X_j)$ *anni/*X_i *con calore atteso pari a* $E(X_i)$ *anni*, indipendentemente dalla durata di corretto funzionamento dell'altra. Quando **almeno una/tutte** delle n è guasta, l'apparecchiatura non è più operativa. Sia T il tempo di corretto funzionamento dell'apparecchiatura. Si esprima T come funzione di X_1, X_2. Si dica qual è il supporto di T. Si ottengano poi la funzione di ripartizione e la funzione di densità di probabilità di T, esplicitandole in tutti i loro tratti. Si calcolino l'n-esimo quartile di T (è il quantile-p con p $= n/100$) e la probabilità condizionale $P(T > b \,|\, T > a)$.

4.2 Soluzione

Bisogna trovare il tempo T di corretto funzionamento dell'apparecchio con n resistenze (solitamente $n = 2$ oppure $n = 3$, supponiamo n abbia quest'ultimo valore) con vita operativa X_i con $i = 1, \dots, n$ e con un certo valore atteso $E(X_i)$.

Bisogna anzitutto esprimere T come funzione delle X_i. Per risolvere questo punto bisogna fare attenzione al testo del problema: se l'apparecchio smette di funzionare quando **almeno** una delle resistenze si guasta allora bisogna utilizzare la funzione **minimo**, se invece l'apparecchio smette di funzionare quando **tutte** le resistenze sono guaste allora bisogna utilizzare la funzione **massimo**. Questo perché il tempo di corretto funzionamento dell'apparecchio è dovuto alle vite operative delle resistenze, quindi con la funzione **minimo** scegliamo il tempo di corretto funzionamento della prima resistenza che si rompe (che soddisfa la richiesta **almeno**) invece con la funzione **massimo** scegliamo il tempo di corretto funzionamento dell'ultima resistenza che si rompe (che soddisfa la richiesta **tutte**). Quindi:

$$T = \begin{cases} \min(X_1, \dots, X_n) & \text{se la richiesta è } \textbf{almeno} \\ \max(X_1, \dots, X_n) & \text{se la richiesta è } \textbf{tutte} \end{cases}$$

Per il compito: *Dire per quale motivo si è scelta la funzione* max *o* min *(vedi testo sopra la formula).*

Dato che le v.c. X_i modellano un tempo d'attesa, allora la loro legge è esponenziale: $X_i \sim Esp(\lambda)$ per $i = 1, \dots, n$. Sappiamo che il supporto di una v.c. con legge esponenziale è $[0, +\infty[$ e che una v.c. trasformata di una v.c. con legge esponenziale avrà anch'essa legge esponenziale, quindi possiamo concludere che:

$$S_T = [0, +\infty[$$

Per il compito: *Indicare esplicitamente che il minimo/massimo di due intervalli tra* $[0, +\infty[$ *è sempre* $[0, +\infty[$

Per trovare il parametro λ basta considerare il valore atteso delle resistenze: se il testo dice *"ciascuna di esse ha distribuzione esponenziale con valore atteso pari a* $E(X_j)$ *anni"* con $j \in [1, n]$ casuale dato che $E(X_i) = E(X_{i+1})$ $\forall x \in [1, n-1]$ si ha che

$$\lambda_i = \frac{1}{E(X_j)}$$

e quindi $\lambda_i = \lambda_j$ per $i = 1, \ldots, n$. Avremo che $\sum_{i=1}^n \lambda_i = \frac{n}{E(X_j)}$.

Se invece il testo elenca i vari valori attesi per le varie X_i (i.e. *ciascuna di esse ha distribuzione esponenziale,* X_1 *con calore atteso pari a* $E(X_1)$ *anni,* $\ldots$*,* X_n *con calore atteso pari a* $E(X_n)$ *anni*) allora avremo che

$$\lambda_i = \frac{1}{E(X_i)}$$

per $i = 1, \ldots, n$. Avremo che $\sum_{i=1}^n \lambda_i = \sum_{i=1}^n \frac{1}{E(X_i)}$.

Il parametro λ ci serve per calcolare la funzione di ripartizione:

$$\begin{aligned}
F_T(t) &= P(T \leq t) = 1 - P(T > t) \\
&= 1 - P(X_1 > t, \ldots, X_n > t) \\
&= 1 - \prod_{i=1}^{n} P(X_i > t) \\
&= 1 - \prod_{i=1}^{n} e^{-\lambda_i t} \\
&= 1 - e^{-t \sum_{i=1}^n \lambda_i}
\end{aligned}$$

ed esplicitandola in tutti i tratti:

$$F_T(t) = \begin{cases} 0 & \text{se } t < 0 \\ 1 - e^{-t \sum_{i=1}^n \lambda_i} & \text{se } t \geq 0 \end{cases}$$

Per il compito: *Indicare assolutamente che* $1 - P(X_1 > t, \ldots, X_n > t) = 1 - \prod_{i=1}^n P(X_i > t)$ *poiché le v.c.* X_i *sono indipendenti.*

La f.d.p. è la derivata della funzione di ripartizione, quindi:

$$\begin{aligned}
p_T(t) &= \frac{d}{dt} F_T(t) \\
&= \frac{d}{dt}\left(1 - e^{-t\sum_{i=1}^n \lambda_i}\right) \\
&= -e^{-t\sum_{i=1}^n \lambda_i} \frac{d}{dt}\left(-t\sum_{i=1}^n \lambda_i\right) \\
&= -e^{-t\sum_{i=1}^n \lambda_i}\left(-\sum_{i=1}^n \lambda_i\right) \\
&= \left(\sum_{i=1}^n \lambda_i\right) e^{-t\sum_{i=1}^n \lambda_i}
\end{aligned}$$

ed esplicitandola in tutti i tratti:

$$p_T(t) = \begin{cases} 0 & \text{se } t < 0 \\ \left(\sum_{i=1}^n \lambda_i\right) e^{-t\sum_{i=1}^n \lambda_i} & \text{se } t \geq 0 \end{cases}$$

Il quantile-p t_p si calcola ponendo la $F_T(t_p) = p$. Dopo dei semplici calcoli si ottiene che:

$$\begin{aligned}
F_T(t_p) &= p \\
1 - e^{-t_p\sum_{i=1}^n \lambda_i} &= p \\
e^{-t_p\sum_{i=1}^n \lambda_i} &= 1 - p \\
-t_p\sum_{i=1}^n \lambda_i &= \log(1-p) \\
t_p &= -\frac{\log(1-p)}{\sum_{i=1}^n \lambda_i}
\end{aligned}$$

Come ultima cosa manca la probabilità condizionata $P(T > b \,|\, T > a)$. Prima di tutto osserviamo che:

$$\begin{aligned}
P(T > t) &= 1 - F_T(t) \\
&= 1 - (1 - e^{-t\sum_{i=1}^n \lambda_i}) \\
&= 1 - 1 + e^{-t\sum_{i=1}^n \lambda_i} \\
&= e^{-t\sum_{i=1}^n \lambda_i}
\end{aligned}$$

Adesso possiamo procedere con la formula classica della probabilità condizionata:

$$\begin{aligned} P(T > b \,|\, T > a) &= \frac{P(T > b)}{P(T > a)} \\ &= \frac{e^{-b\sum_{i=1}^{n}\lambda_i}}{e^{-a\sum_{i=1}^{n}\lambda_i}} \\ &= e^{-b\sum_{i=1}^{n}\lambda_i + a\sum_{i=1}^{n}\lambda_i} \\ &= e^{(a-b)\sum_{i=1}^{n}\lambda_i} \\ &= e^{-(b-a)\sum_{i=1}^{n}\lambda_i} \end{aligned}$$

5 Funzione generatrice dei momenti

5.1 Testo

Sia Y una variabile casuale univariata avente quale funzione generatrice dei momenti $M_Y(t) = exp(g(t))$, dove $\exp(z) = e^z$. Siano poi $Y_1, \dots, Y_n$ copie indipendenti di Y e si ponga $W = c_1 * Y_1 + \cdots + c_n * Y_n$. Si calcoli la funzione generatrice dei momenti di W. Si ottengano valore atteso e varianza di W.

5.2 Soluzione

Abbiamo la variabile Y con funzione generatrice dei momenti $M_Y(t) = e^{g(t)}$ e $Y_1, \dots, Y_n$ copie indipendenti di Y, che vuol dire che $Y \sim Y_1 \sim \cdots \sim Y_n$. Abbiamo poi W definito come funzione delle Y_i per $i = 1, \dots, n$, ma a noi conviene definire W anche come $W = \sum_{i=1}^n c_i Y_i = c_1 Y_1 + \cdots + c_n Y_n$. In questo modo possiamo vedere $W = -Y_1 - Y_2$ come $W = c_1 Y_1 + c_2 Y_2$ con $c_1(x) = c_2(x) = -1$.

La funzione generatrice dei momenti si ottiene tramite i seguenti passaggi:

$$
\begin{aligned}
M_W(t) &= E(e^{tW}) \\
&= E(e^{t \sum_{i=1}^n c_i Y_i}) \\
&= E\left(\prod_{i=1}^n e^{tc_i Y_i}\right) \\
&= \prod_{i=1}^n E(e^{c_i t Y_i}) \\
&= \prod_{i=1}^n M_{Y_i}(c_i t) \\
&= \prod_{i=1}^n M_Y(c_i t) \\
&= \prod_{i=1}^n e^{g(c_i t)} \\
&= e^{\sum_{i=1}^n g(c_i t)}
\end{aligned}
$$

ed eventualmente si semplifica.

Per il compito: *Indicare assolutamente che* $E\left(\prod_{i=1}^n e^{tc_i Y_i}\right) = \prod_{i=1}^n E(e^{c_i t Y_i})$ *per l'indipendenza delle* Y_i.

Per il compito: *Conviene indicare che* $\prod_{i=1}^n M_{Y_i}(c_i t) = \prod_{i=1}^n M_Y(c_i t)$ *poiché* $Y \sim Y_1 \sim \cdots \sim Y_n$.

Per il valore atteso bisogna calcolare la derivata $M'_W(t)$. Si avrà che:

$$E(X) = M'_W(0)$$

Per la varianza conviene calcolare la derivata seconda $M''_W(t)$. Si avrà che:

$$E(X^2) = M''_W(0)$$

e poi procedere con la solita formula

$$Var(X) = E(X^2) - E(X)^2$$

6 Leggi normali

6.1 Testo

La variabile casuale multivariata $(Y_1, \ldots, Y_n)$ ha componenti indipendenti e identicamente distribuite con legge marginale normale, in particolare $Y_1 \sim N(\mu, \sigma^2)$. Si mostri che la variabile casuale $\overline{Y}_n = \sum_{i=1}^{n} \frac{Y_i}{n}$ ha legge normale, $\overline{Y}_n \sim N(\mu, \sigma^2/n)$. Sia $n = n_0$. Si calcolino $P(\overline{Y}_{n_0} > a)$ e $P(\overline{Y}_{n_0} < b)$. Si ottenga infine l'$i$-esimo percentile di $\overline{Y}_{n_0}$ (il quantile-p con $p = 0.i$).

6.2 Soluzione

Dobbiamo dimostrare che la v.c. $\overline{Y}_n = \sum_{i=1}^{n} \frac{Y_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ sapendo che $Y_1 \sim N(\mu, \sigma^2)$ e che $Y_1 \sim \cdots \sim Y_n$.

Si ricorda che se $Y \sim N(\mu, \sigma^2)$ allora avrà funzione generatrice dei momenti $M_Y(t) = E(e^{t\mu + \frac{1}{2}t^2\sigma^2})$.

La dimostrazione è schematica e richiede i seguenti passaggi:

$$\begin{aligned}
M_{\overline{Y}_n}(t) &= E\left(e^{t\overline{Y}_n}\right) \\
&= E\left(e^{t\frac{1}{n}\sum_{i=1}^{n} Y_i}\right) \\
&= E\left(\prod_{i=1}^{n} e^{\frac{t}{n}Y_i}\right) \\
&= \prod_{i=1}^{n} E\left(e^{\frac{t}{n}Y_i}\right) \\
&= \left(E\left(e^{\frac{t}{n}Y_1}\right)\right)^n \\
&= \left(M_{Y_1}\left(\frac{t}{n}\right)\right)^n \\
&= \left(e^{\frac{t}{n}\mu + \frac{1}{2}\left(\frac{t}{n}\right)^2\sigma^2}\right)^n \\
&= e^{n\left(\frac{t}{n}\mu + \frac{1}{2}\frac{t^2}{n^2}\sigma^2\right)} \\
&= e^{t\mu + \frac{1}{2}t^2\frac{\sigma^2}{n}}
\end{aligned}$$

da cui si conclude che $\overline{Y}_n = \sum_{i=1}^{n} \frac{Y_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ poiché una funzione generatrice dei momenti (in questo caso $M_{\overline{Y}_n}(t)$) caratterizza la legge di probabilità della v.c. associata (in questo caso $\overline{Y}_n$).

Per il compito: *Indicare assolutamente che* $E\left(\prod_{i=1}^{n} e^{\frac{t}{n}Y_i}\right) = \prod_{i=1}^{n} E\left(e^{\frac{t}{n}Y_i}\right)$ *per l'indipendenza delle* Y_i.

Per il compito: *Indicare assolutamente che* $\prod_{i=1}^{n} E\left(e^{\frac{t}{n}Y_i}\right) = \left(E\left(e^{\frac{t}{n}Y_1}\right)\right)^n$ *per l'identica distribuzione delle* Y_i.

Per il calcolo delle varie probabilità si ricorda che $P(\overline{Y}_n > t) = 1 - P(\overline{Y}_n \leq t)$ e che $P(\overline{Y}_n \leq t) = P(\overline{Y}_n < t)$ poiché $P(\overline{Y}_n = t) = 0$ se la v.c. ha legge continua.

A questo punto ci basta sapere come calcolare $P(\overline{Y}_n \leq t) = F_{\overline{Y}_n}(t)$

$$F_{\overline{Y}_n}(t) = \Phi\left(\frac{t-\mu}{\sqrt{\frac{\sigma^2}{n}}}\right) = \Phi\left(\frac{t-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = \Phi\left(\sqrt{n}\,\frac{t-\mu}{\sigma}\right)$$

ATTENZIONE: sigma non è elevato al quadrato!

Se $t - \mu < 0$ si ricorre alla formula $\Phi(-x) = 1 - \Phi(x)$ con $-x = \sqrt{n}\,\frac{t-\mu}{\sigma}$.

Per ottenere il valore di $\Phi(x)$ bisogna fare riferimento alle tavole della funzione di ripartizione della legge normale standard.

Per il quantile-p t_p anche qui bisogna porre $F_{\overline{Y}_n}(t) = p$ passando però tramite la v.c. $Z \sim N(0,1)$. Sia quindi

$$t_p = \mu + \sqrt{\frac{\sigma^2}{n}} z_p = \mu + \frac{\sigma}{\sqrt{n}} z_p$$

si ottiene che

$$F_{\overline{Y}_n}(t) = \Phi\left(\frac{t-\mu}{\sqrt{\frac{\sigma^2}{n}}}\right) = \Phi\left(\frac{\mu + \sqrt{\frac{\sigma^2}{n}} z_p - \mu}{\sqrt{\frac{\sigma^2}{n}}}\right) = \Phi(z_p) = p$$

da cui possiamo ricavare

$$z_p = \Phi^{-1}(p)$$

Nelle tavole è riportato solo $p \geq 0,5$, ma questo non è un problema tenendo conto della formula $z_p = -z_{1-p}$ e quindi

$$\Phi^{-1}(p) = -\Phi^{-1}(1-p)$$

In conclusione quello che ci basta fare per calcolare un quantile-p è usare la formula

$$t_p = \mu + \sqrt{\frac{\sigma^2}{n}}\,\Phi^{-1}(p) = \mu + \frac{\sigma}{\sqrt{n}}\,\Phi^{-1}(p)$$

ed eventualmente applicare la formula precedente se $p < 0,5$.

Nota: il quantile-p con $p = 0,5$, ovvero la mediana, è sempre pari a 0 per le v.c. con leggi normali.

6.2.1 Calcolare Φ e Φ^{-1}

La tavola M della funzione Φ si presenta sotto forma di matrice: nell'intestazione delle righe (quindi nella prima colonna M^1) sono riportate le prime due cifre t_c di t (la cifra intera e la prima cifra decimale), mentre nell'intestazione delle colonne (quindi nella prima riga M_1) è riportata l'ultima cifra decimale t_r.

Per calcolare $\Phi(t)$ dobbiamo cercare $t_c + t_r = t$ con $t_c = m_{i,1}$ e $t_r = m_{1,j}$. Il valore $\Phi(t) = m_{i,j}$, dove $m_{i,j}$ è l'elemento all'incrocio tra la colonna di t_r e la riga di t_c.

Per calcolare $\Phi^{-1}(p)$ basta trovare l'elemento $m_{i,j} = p$, in caso approssimando p al valore più vicino presente nella tabella e guardare gli elementi nelle intestazioni $m_{i,1}$ e $m_{1,j}$. Anche qui avremo che $t = t_c + t_r$ con $t_c = m_{i,1}$ e $t_r = m_{1,j}$.

7 Popolazioni e stime

7.1 Testo

Dato un campione $y_1, \ldots, y_n$, realizzazione di variabili casuali $Y_1, \ldots, Y_n$, $n > 1$, indipendenti con legge esponenziale con tasso di guasto $\lambda > 0$ ignoto/di Poisson con valore atteso $\lambda > 0$ ignoto/normale di media nota pari a μ e varianza $\sigma^2 > 0$ ignota/normale di media $\mu \in \mathbb{R}$ ignota e varianza nota pari a σ^2/$Bi(n,p)$ di media $p \in (a,b)$ ignota, si reperisca una stima di $\lambda/\lambda/\sigma^2/\mu/p$ e si indaghino le proprietà campionarie dello stimatore corrispondente.

calcolandone anche lo standard error (solo nel caso della legge $Bi(n,p)$).

7.2 Soluzioni

7.2.1 Richiami: la v.c. media campionaria

Sia $Y = (Y_1, \ldots, Y_n)$, $E(Y_i) = \mu$ e $Var(Y_i) = \sigma^2$ per ogni $i = 1, \ldots, n$. Si ricorda che $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$ ha valore atteso

$$E(\overline{Y}_n) = E\left(\frac{1}{n}\sum_{i=1}^n Y_i\right) = \frac{1}{n}\sum_{i=1}^n E(Y_i) = \frac{1}{n}\sum_{i=1}^n \mu = \frac{1}{n}n\mu = \mu$$

e varianza

$$Var(\overline{Y}_n) = Var\left(\frac{1}{n}\sum_{i=1}^n Y_i\right) = \frac{1}{n^2}\sum_{i=1}^n Var(Y_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

Si ricorda inoltre che se $y = (y_1, \ldots, y_n)$ è una realizzazione di Y, allora la realizzazione $\overline{y}_n$ di $\overline{Y}_n$ è

$$\overline{y}_n = \frac{1}{n}\sum_{i=1}^n y_i$$

7.2.2 Poisson

Sia $Y = (Y_1, \ldots, Y_n)$ una variabile di c.c.s. con numerosità n e legge $Y \sim P(\lambda)$ (Poisson), allora una stima $\hat{\lambda}_n$ per λ è:

$$\hat{\lambda}_n(y) = \overline{y}_n$$

Stima	$\hat{\lambda}_n(y) = \overline{y}_n$
Stimatore	$\hat{\lambda}_n = \overline{Y}_n$
Valore atteso	$E(\hat{\lambda}_n) = \lambda$
Varianza	$Var(\hat{\lambda}_n) = \frac{\lambda}{n}$
Standard error	$\hat{se}(\hat{\lambda}_n) = \frac{\sqrt{\hat{\lambda}_n}}{\sqrt{n}}$

7.2.3 Esponenziale

Sia $Y = (Y_1, \dots, Y_n)$ una variabile di c.c.s. con numerosità n e legge $Y \sim Esp(\lambda)$, allora una stima $\hat{\lambda}_n$ per λ è:

$$\hat{\lambda}_n(y) = \frac{1}{\overline{y}_n}$$

Stima	$\hat{\lambda}_n(y) = \frac{1}{\overline{y}_n}$
Stimatore	$\hat{\lambda}_n = \frac{1}{\overline{Y}_n}$
Valore atteso	$E(\hat{\lambda}_n) = \frac{1}{\lambda}$
Varianza	$Var(\hat{\lambda}_n) = \frac{\frac{1}{\lambda^2}}{n} = \frac{1}{n\lambda^2}$
Standard error	$\hat{se}(\hat{\lambda}_n) = \frac{1}{\sqrt{n\hat{\lambda}_n^2}} = \frac{1}{\hat{\lambda}_n\sqrt{n}}$

7.2.4 Normale con μ noto

Sia $Y = (Y_1, \dots, Y_n)$ una variabile di c.c.s. con numerosità n e legge $Y \sim N(\mu, \sigma^2)$ con μ noto, allora una stima $\hat{s}_n^2$ per σ^2 è:

$$\hat{\overline{\mu}}_2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2$$

Stima	$\hat{\overline{\mu}}_2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2$
Stimatore	$\hat{\overline{\mu}}_2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu)^2$
Valore atteso	$E(\hat{\overline{\mu}}_2) = \sigma^2$
Varianza	$Var(\hat{\overline{\mu}}_2) = \frac{2\sigma^4}{n}$
Standard error	$\hat{se}(\hat{\overline{\mu}}_2) = \frac{\sqrt{2\hat{\overline{\mu}}_2^2}}{\sqrt{n}} = \frac{\hat{\overline{\mu}}_2\sqrt{2}}{\sqrt{n}}$

7.2.5 Normale con σ^2 noto

Sia $Y = (Y_1, \dots, Y_n)$ una variabile di c.c.s. con numerosità n e legge $Y \sim N(\mu, \sigma^2)$ con σ^2 noto, allora una stima $\hat{\mu}_n$ per μ è:

$$\hat{\mu}_n(y) = \overline{y}_n$$

Stima	$\hat{\mu}_n(y) = \overline{y}_n$
Stimatore	$\hat{\mu}_n = \overline{Y}_n$
Valore atteso	$E(\hat{\mu}_n) = \mu$
Varianza	$Var(\hat{\mu}_n) = \frac{\sigma^2}{n}$
Standard error	$\hat{se}(\hat{\mu}_n) = \frac{\sqrt{\sigma^2}}{\sqrt{n}}$

7.2.6 Normale con μ e σ^2 ignoti

Sia $Y = (Y_1, \ldots, Y_n)$ una variabile di c.c.s. con numerosità n e legge $Y \sim N(\mu, \sigma^2)$ con μ e σ^2 ignoti, allora una stima $\hat{s}_n^2$ per σ^2 è:

$$\hat{s}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \mu)^2$$

Stima	$\hat{s}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \mu)^2$
Stimatore	$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$
Valore atteso	$E(\hat{s}_n^2) = \sigma^2$
Varianza	$Var(\hat{\sigma}_n) = \frac{2\sigma^4}{n-1}$
Standard error	$\hat{se}(\hat{s}_n^2) = \frac{\sqrt{2\hat{s}_n^4}}{\sqrt{n-1}} = \frac{\hat{s}_n^2\sqrt{2}}{\sqrt{n-1}}$

7.2.7 Binomiale

Sia $Y = (Y_1, \ldots, Y_n)$ una variabile di c.c.s. con numerosità n e legge $Y \sim Bi(n, p)$ con $p \in \,]a, b[$ e n noto, allora una stima $\hat{p}_n$ per p è:

$$\hat{p}_n(y) = \overline{y}_n$$

Stima	$\hat{p}_n(y) = \overline{y}_n$
Stimatore	$\hat{p}_n = \overline{Y}_n$
Valore atteso	$E(\hat{p}_n) = p$
Varianza	$Var(\hat{p}_n) = \frac{p(1-p)}{n}$
Standard error	$\hat{se}(\hat{p}_n) = \frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}$

7.3 Verificare le proprietà

Tutti gli stimatori riportati in questa dispensa godono di tutte le proprietà elencate.

7.3.1 Distorsione

Uno stimatore $\hat{\theta}_n$ è detto **non distorto** per θ se

$$E(\hat{\theta}_n) = \theta$$

Questa proprietà si verifica facilmente se lo stimatore è la v.c. $\overline{Y}_n$ (vedi inizio sezione), altrimenti bisogna utilizzare le varie proprietà del valore atteso: si parte dal termine $E(\hat{\theta}_n)$ e si cerca di ottenere θ.

7.3.2 Distorsione asintotica

Uno stimatore $\hat{\theta}_n$ è detto **asintoticamente non distorto** per θ se

$$\lim_{n \to +\infty} E(\hat{\theta}_n) = \theta$$

Questa proprietà si verifica facilmente se $E(\hat{\theta}_n) = \theta$ (ovvero lo stimatore è non distorto): dato che θ non dipende da n se lo stimatore è non distorto, allora $E(\hat{\theta}_n) = \theta \;\; \forall n \in \mathbb{N}$.

7.3.3 Consistenza

Uno stimatore $\hat{\theta}_n$ è detto **consistente** per θ se

$$\hat{\theta}_n \xrightarrow{p} \theta$$

Sia $Y = (Y_1, \ldots, Y_n)$ una v.c. multivariata con componenti Y_i indipendenti e identicamente distribuite con valore atteso $E(Y_1) = \mu$ e varianza $Var(Y_1) = \sigma^2$ finite (si veda le tabelle dei vari stimatori).

Per la *legge dei grandi numeri* $\overline{Y}_n \xrightarrow{p} \mu$ con $\overline{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i$ la media campionaria. Dato che $\mu = E(Y_1) = E(\overline{Y}_n)$ e che la media campionaria ha come valore atteso la stima θ, $\overline{Y}_n \xrightarrow{p} \theta$. Quindi uno stimatore $\hat{\theta}_n$ per θ in Y, se deriva dalla media campionaria è consistente.

7.3.4 Normalità asintotica

Uno stimatore $\hat{\theta}_n$ è detto **asintoticamente normale** per θ se

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$$

dove $\sigma^2(\theta)$ è la varianza asintotica.

Si può dimostrare usando il *teorema del limite centrale* che $\overline{Y}_n$ è uno stimatore asintoticamente normale.

8 Popolazioni e stime (Alternativo)

8.1 Testo

Il responsabile qualità di una fabbrica vuole valutare la proporzione p di pezzi prodotti che non sono idonei alla vendita. Viene estratto dalla produzione del giorno un campione casuale di n pezzi e si osserva che d di essi non sono idonei. Si definisca un modello statistico parametrico utile per descrivere l'esperimento effettuato. Si reperisca uno stimatore per p, si determini l'associato standard error e si calcolino le corrispondenti stime.

8.2 Soluzione

Vogliamo valutare la porzione p sapendo che n sono i pezzi totali considerati e d sono i pezzi difettosi. Il modello statistico è quindi

$$Y = (Y_0, \dots, Y_n) \text{ con } Y_1 \sim \dots \sim Y_n \sim Bi(1, p)$$

Le componenti Y_i di Y saranno quindi indipendenti e identicamente distribuite, come al solito.

Avendo $y = (y_1, \dots, y_n)$ realizzazione di Y, il parametro d equivale al numero di pezzi difettosi, quindi

$$d = \sum_{i=1}^{n} y_i$$

dato che $y \in \mathcal{Y} = \{0, 1\}^n$ (poiché $Y_i \sim Bi(1, n)$ quindi $S_{Y_i} = \{0, 1\}$) e che $y_i = 0$ se il pezzo i NON è difettoso, $y_i = 1$ se il pezzo i è difettoso.

Lo stimatore $\hat{p}_n$ per p è quindi (come già riportato)

$$\hat{p}_n = \overline{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

La stima invece è

$$\hat{p}_n(y) = \overline{y}_n = \frac{1}{n} \sum_{i=1}^{n} y_i$$

dove $y = (y_1, \dots, y_n)$ è la realizzazione di Y.

Sapendo che $\sum_{i=1}^{n} y_i = d$ abbiamo che il valore della stima $\hat{p}_n(y)$ è

$$\hat{p}_n(y) = \overline{y}_n = \frac{1}{n} \cdot d = \frac{d}{n}$$

Questo vuol dire che la porzione di oggetti difettosi $p \simeq \frac{d}{n}$, come ci si può aspettare logicamente.

(Necessario?) Il valore atteso di $\overline{Y}_n$ è

$$E(\overline{Y}_n) = E\left(\frac{S_n}{n}\right) = \frac{1}{n}E(S_n) = \frac{1}{n}np = p$$

con $S_n = \sum_{i=1}^{n} Y_i$.

(Necessario?) La varianza di $\overline{Y}_n$ è

$$Var(\overline{Y}_n) = Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2}Var(S_n) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}$$

ed è maggiorata da $\frac{1}{4n}$ (i.e. $Var(\overline{Y}_n) \leq \frac{1}{4n}$). Si ha sempre $S_n = \sum_{i=1}^{n} Y_i$.

Infine lo standard error $\hat{se}(\hat{p}_n)$ è

$$\hat{se}(\hat{p}_n) = \frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}$$

poiché $\hat{se}(\hat{p}_n) = \sqrt{Var(\hat{p}_n)}$ dato che $\overline{Y}_n$ è uno stimatore non distorto (i.e. $E(\overline{Y}_n) = \hat{p}_n$).

(Per completezza) La probabilità di una realizzazione y_i, di una singola componente $Y_i \sim Bi(1,p)$, di essere difettosa è:

$$p_Y(y_i;p) = \binom{1}{y_i} p^{y_i}(1-p)^{1-y_i} = p^{y_i}(1-p)^{1-y_i}$$

poiché $y_i \in 0,1$ e quindi $\binom{1}{y_i} = 1$ per ogni valore di y_i.

Da questo segue che la probabilità $p_Y(y;p)$ della realizzazione y di Y di avere d pezzi difettosi è

$$\begin{aligned} p_Y(y;p) &= \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i} \\ &= p^{\sum_{i=1}^{n} y_i}(1-p)^{n-\sum_{i=1}^{n} y_i} \\ &= p^d(1-p)^{n-d} \end{aligned}$$