Esame di Calcolo delle Probabilità

(2 ore e 30 minuti)

Si prega di scrivere in maniera chiara, risposte non leggibili non saranno corrette. In tutti gli esercizi si richiede di illustrare il proprio lavoro.

Esercizio 1 (10 punti). Si consideri un gruppo di $n = 3650$ coppie di sposi. Assumiamo che ciascuna persona, indipendentemente da ciascun altra, sia nata in un giorno casuale (cioè la probabilità dell'evento {La persona è nata il giorno a} è $1/365$, per ogni giorno a dell'anno) di un anno non bisestile.

(a) Consideriamo una coppia fissata: qual è la probabilità p che i coniugi siano nati lo stesso giorno? Qual è la probabilità q che i coniugi siano nati lo stesso mese? [Non è necessario semplificare le frazioni ottenute.]

(b) Si scriva un codice R per fornire una stima del valore di p.

Sia X la variabile aleatoria che conta il numero di coppie nate lo stesso giorno e sia Y la variabile aleatoria che conta il numero di coppie nate lo stesso mese. [Le prossime risposte si possono fornire in termini di n, p, e q senza esplicitare i numeri.]

(c) Si determinino valore medio e varianza delle variabili aleatorie X e Y.

(d) Le variabili X e Y sono indipendenti?

(e) Si calcoli la varianza di $Z = Y - X$.

Soluzione. (a) Siano U e V le variabili aleatorie che contengono il giorno in cui è nato il primo e il secondo coniuge, rispettivamente. Si ha

$$\mathbb{P}(U = a) = \mathbb{P}(V = a) = \frac{1}{365}.$$

Vogliamo calcolare $\mathbb{P}(U = V)$. L'evento $\{U = V\}$ si può scrivere come l'unione disgiunta

$$\{U = V\} = \bigcup_{a=1}^{365} \{U = a, V = a\}.$$

Usando il fatto che U e V sono indipendenti si ha

$$p = \mathbb{P}(U = V) = \sum_{a=1}^{365} \mathbb{P}(U = a, V = a) = \sum_{a=1}^{365} \mathbb{P}(U = a)\mathbb{P}(V = a) = \sum_{a=1}^{365} \frac{1}{365^2} = \frac{1}{365}.$$

Consideriamo l'evento $M = \{$I coniugi sono nati lo stesso mese$\}$ e definiamo gli eventi

$$A = \{\text{Il coniuge 1 è nato in un mese di 31 giorni}\},$$
$$B = \{\text{Il coniuge 1 è nato in un mese di 30 giorni}\},$$
$$C = \{\text{Il coniuge 1 è nato in un mese di 28 giorni}\}.$$

Sappiamo che $\mathbb{P}(M) = \mathbb{P}(M|A)\mathbb{P}(A) + \mathbb{P}(M|B)\mathbb{P}(B) + \mathbb{P}(M|C)\mathbb{P}(C)$. Nell'anno ci sono 7 mesi di 31 giorni, 4 mesi di 30 giorni, e 1 mese di 28 giorni, quindi

$$\mathbb{P}(A) = \frac{7 \cdot 31}{365}, \quad \mathbb{P}(B) = \frac{4 \cdot 30}{365}, \quad \mathbb{P}(C) = \frac{28}{365}.$$

D'altra parte $\mathbb{P}(M|A)$ è la probabilità che il secondo coniuge sia nato in un certo mese fissato di 31 giorni, quindi

$$\mathbb{P}(M|A) = \frac{31}{365},$$

similmente

$$\mathbb{P}(M|B) = \frac{30}{365} \quad \text{e} \quad \mathbb{P}(M|C) = \frac{28}{365}.$$

Quindi abbiamo

$$q = \frac{7 \cdot 31^2}{365^2} + \frac{4 \cdot 30^2}{365^2} + \frac{28^2}{365^2}.$$

(b) Possiamo utilizzare il seguente codice

```
giorni=c(1:365);
  n=100000;
  s=0;
  for(i in c(1:n)){
    compleanno_c1=sample(giorni,size=1);
    compleanno_c2=sample(giorni,size=1);
    if (compleanno_c1==compleanno_c2){
      s=s+1;
  }
  p_empirica=s/nG
```

(c) Ovviamente $X \sim B(n,p)$ e $Y \sim B(n,q)$, quindi

$$\mathbb{E}(X) = np, \quad Var(X) = np(1-p),$$
$$\mathbb{E}(Y) = nq, \quad Var(Y) = nq(1-q).$$

(d) Le due variabili non sono indipendenti. Infatti,

$$\mathbb{P}(X = 1, Y = 0) = 0,$$

ma

$$\mathbb{P}(X = 1) \neq 0 \quad \text{e} \quad \mathbb{P}(Y = 0) \neq 0,$$

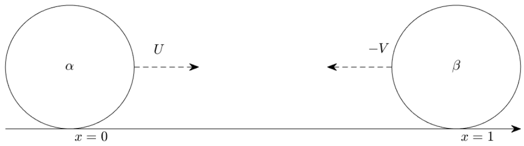
allora

$$\mathbb{P}(X = 1, Y = 0) \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 0).$$

(e) La variabile Z conta il numero di coppie che sono nate nello stesso mese, ma in giorni differenti. La probabilità che due persone siano nate nello stesso mese, ma in giorni differenti è $r = q - p$. Allora $Z \sim B(n,r)$, da cui

$$Var(Z) = nr(1-r) = n(q-p)(1-q+p).$$

Esercizio 2 (6 punti). Due particelle puntiformi, che indichiamo con α e β, si muovono lungo una retta, entrambe di moto rettilineo uniforme. All'istante iniziale $t = 0$ la particella α si trova nell'origine $x = 0$ e la particella β si trova in $x = 1$. Le due particelle si muovono l'una verso l'altra, con velocità rispettive U e $-V$, dove U e V sono variabili aleatorie indipendenti con distribuzione $Esp(1)$.



Definiamo T come l'istante in cui le due particelle si incontrano, e poniamo

$$Y = \frac{1}{T}.$$

(a) Si mostri che $Y \sim \Gamma(2, 1)$.[1]

(b) Si deduca che T è assolutamente continua con densità

$$f_T(t) = \frac{1}{t^3} \mathrm{e}^{-1/t} \mathbf{1}_{(0,\infty)}(t).$$

(c) Si dimostri che $T \in L^1$. Si mostri che $Cov(T, Y)$ è ben definita, e la si calcoli.

Soluzione. (a) La posizione al tempo t di α è Ut e la posizione al tempo t di β è $1 - Vt$. Il tempo T è definito come quell'istante tale che

$$UT = 1 - VT,$$

Quindi abbiamo

$$T = \frac{1}{U + V}$$

e di conseguenza

$$Y = U + V.$$

Poiché $U \sim Esp(1) = \Gamma(1, 1)$ e $V \sim Esp(1) = \Gamma(1, 1)$ sono indipendenti, abbiamo che

$$Y \sim \Gamma(1 + 1, 1) = \Gamma(2, 1).$$

(b) Calcoliamo F_T, la funzione di ripartizione di T. Ovviamente per $t \leq 0$ abbiamo $F_T(t) = 0$. Sia $t > 0$ e denotiamo con f_Y la densità di Y, allora

$$\begin{aligned}
F_T(t) &= \mathbb{P}(T \leq t) = \mathbb{P}\left(\frac{1}{Y} \leq t\right) = \mathbb{P}\left(Y \geq \frac{1}{t}\right) \\
&= \int_{1/t}^{\infty} f_Y(y) dy = \int_{1/t}^{\infty} y \mathrm{e}^{-y} dy = -y\mathrm{e}^{-y}|_{1/t}^{\infty} + \int_{1/t}^{\infty} \mathrm{e}^{-y} dy \\
&= 0 + \frac{1}{t}\mathrm{e}^{-1/t} + \mathrm{e}^{-1/t} = \mathrm{e}^{-1/t}\left(1 + \frac{1}{t}\right).
\end{aligned}$$

[1] Si ricordi che, se $X \sim \Gamma(\alpha, \lambda)$ e $Y \sim \Gamma(\beta, \lambda)$ sono indipendenti, allora $X + Y \sim \Gamma(\alpha + \beta, \lambda)$

Poiché F_T è continua e quasi ovunque differenziabile T è assolutamente continua e una sua densità si può ottenere differenziando F_T.

Per $t \leq 0$ si ha $f_T(t) = 0$, mentre per $t > 0$

$$f_T(t) = \frac{1}{t^2}\mathrm{e}^{-1/t}\left(1 + \frac{1}{t}\right) - \mathrm{e}^{-1/t}\frac{1}{t^2} = \frac{1}{t^3}\mathrm{e}^{-1/t}.$$

(c) Poiché $T \geq 0$ quasi certamente $\mathbb{E}(T) = \mathbb{E}(|T|)$, quindi ci basta calcolare $\mathbb{E}(T)$ e verificare che è finito.

$$\mathbb{E}(T) = \int_{-\infty}^{\infty} t f_T(t)dt = \int_0^{\infty} \frac{1}{t^2}\mathrm{e}^{-1/t}dt = \int_{-\infty}^{0} \mathrm{e}^y dy = 1.$$

Poiché $YT \equiv 1$, ovviamente $YT \in L^1$, e $Y \in L^1$ poiché è una Gamma, allora la covarianza è ben definita e vale

$$Cov(T, Y) = \mathbb{E}(TY) - \mathbb{E}(T)\mathbb{E}(Y) = 1 - 1 \cdot 2 = -1.$$

Esercizio 3 (8 punti). Sia (X, Y) un vettore aleatorio assolutamente continuo, con densità

$$f_{X,Y}(x, y) = \frac{1}{x + y}\mathbf{1}_D(x, y), \quad \text{dove} \quad D = \{(x, y) \in \mathbb{R}^2 :\ x > 0,\ y > 0,\ x + y < 1\}.$$

(a) Si determini la distribuzione della variabile aleatoria $Z = X + Y$, riconoscendola come notevole.

(b) Si mostri che le componenti X e Y hanno la stessa densità

$$f_X(x) = \log\left(\frac{1}{x}\right)\mathbf{1}_{(0,1)}(x).$$

(c) Si mostri che la funzione di ripartizione $F = F_X$ di X soddisfa la relazione

$$F(x) \sim x \log\left(\frac{1}{x}\right) \quad \text{per} \quad x \to 0^+,$$

cioè che $\lim_{x\to 0^+} \frac{F(x)}{x\log\left(\frac{1}{x}\right)} = 1$.

(d) Siano ora $\{X_n\}$ variabili i.i.d. con la stessa legge di X. Definiamo per $n = 1, 2, \ldots$ la variabile aleatoria reale

$$W_n = n\log(n)\min\{X_1, \ldots, X_n\}.$$

Si mostri che W_n converge in legge per $n \to \infty$ verso una variabile aleatoria W di legge nota.

Soluzione. (a) La densità di Z si può calcolare come

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)dx.$$

Ovviamente $f_Z(z) = 0$ se $z \leq 0$ e se $z \geq 1$. Sia quindi $0 < z < 1$

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{x+z-x} \mathbf{1}_D(x, z-x)dx = \int_{-\infty}^{\infty} \frac{1}{z} \mathbf{1}_D(x, z-x)dx \\ &= \frac{1}{z} \int_0^z dx = 1, \end{aligned}$$

dove la penultima uguaglianza è dovuta al fatto che $\mathbf{1}_D(x, z-x) = 1$ solo se $x > 0$ e $z - x > 0$, cioè, se $x < z$. Abbiamo ottenuto che

$$f_Z(z) = \mathbf{1}_{(0,1)}(z),$$

quindi $Z \sim U((0,1))$.

(b) Per simmetria ovviamente X e Y hanno la stessa legge. Calcoliamo la densità di X. Ovviamente $f_X(x) = 0$ per $x \leq 0$ o $x \geq 1$. Sia $0 < x < 1$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy = \int_0^{1-x} \frac{1}{x+y} dy = \int_x^1 \frac{1}{t} dt = \log(1) - \log(x) = \log\left(\frac{1}{x}\right)$$

(c) Calcoliamo la funzione di ripartizione di X. Ovviamente se $t \leq 0$, allora $F_X(t) = 0$, e se $t \geq 1$, allora $F_X(t) = 1$. Sia $0 < t < 1$

$$F_X(t) = \int_{-\infty}^{t} f_X(x)dx = \int_0^t \log\left(\frac{1}{x}\right) dx = -\int_0^t \log(x)dx = -t\log(t) + t.$$

Calcoliamo ora il limite

$$\lim_{x \to 0^+} \frac{-x\log(x) + x}{x \log\left(\frac{1}{x}\right)} = \lim_{x \to 0^+} \frac{-x\log(x) + x}{-x\log(x)} = \lim_{x \to 0^+} 1 - \frac{1}{\log(x)} = 1.$$

(d) Calcoliamo $F_{W_n}(t)$ la funzione di ripartizione di W_n. Ovviamente $F_{W_n}(t) = 0$ per $t \leq 0$, sia quindi $t > 0$

$$\begin{aligned} F_{W_n}(t) &= \mathbb{P}(W_n \leq t) = \mathbb{P}(n\log(n)\min\{X_1, \ldots, X_n\} \leq t) = \mathbb{P}\left(\min\{X_1, \ldots, X_n\} \leq \frac{t}{n\log(n)}\right) \\ &= 1 - \mathbb{P}\left(\min\{X_1, \ldots, X_n\} > \frac{t}{n\log(n)}\right) = 1 - \mathbb{P}\left(X_1 > \frac{t}{n\log(n)}, \ldots, X_n > \frac{t}{n\log(n)}\right) \\ &= 1 - \left(\mathbb{P}\left(X_1 > \frac{t}{n\log(n)}\right)\right)^n = 1 - \left(1 - F_X\left(\frac{t}{n\log(n)}\right)\right)^n. \end{aligned}$$

Facciamo ora tendere n all'infinito. Per $t \geq 0$ ovviamente $\lim_{n\to\infty} F_{W_n}(t) = 0$. Sia $t > 0$, per il punto precedente abbiamo $F_X\left(\frac{t}{n\log(n)}\right) \sim \frac{t}{n\log(n)} \log\left(\frac{n\log(n)}{t}\right)$ per $n \to \infty$, quindi

$$\begin{aligned}
\lim_{n\to\infty} F_{W_n}(t) &= \lim_{n\to\infty} 1 - \left(1 - \frac{t}{n\log(n)} \log\left(\frac{n\log(n)}{t}\right)\right)^n \\
&= \lim_{n\to\infty} 1 - \left(1 - \frac{t\left(\log(n) + \log(\log(n)) - \log(t)\right)}{n\log(n)}\right)^n \\
&= \lim_{n\to\infty} 1 - \left(1 - \frac{t}{n} + \frac{t\log(\log(n))}{n\log(n)} - \frac{\log(t)}{n\log(n)}\right)^n \\
&= \lim_{n\to\infty} 1 - \left(1 - \frac{t}{n}\right)^n = 1 - \mathrm{e}^{-t}.
\end{aligned}$$

Quindi $W_n \xrightarrow{\mathcal{L}} W$ con $W \sim Esp(1)$.

Tabella 1: $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{t^2/2}\mathrm{d}t$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.508	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
0.1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Tabella 2: Legge, funzione di ripartizione, media, e varianza più importanti distribuzioni.

	Densità	F. di rip.	Media	Varianza
$B(n,p)$	$p_k = \binom{n}{k} p^k (1-p)^{n-k}$		np	$np(1-p)$
Ipergeometrica	$p_k = \frac{\binom{r}{k}\binom{b}{n-k}}{\binom{b+r}{n}}$		$\frac{nr}{b+r}$	$\frac{nrb(b+r-n)}{(b+r)^2(b+r-1)}$
$P(\lambda)$	$p_k = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$		λ	λ
$U([a,b])$	$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \text{altrimenti} \end{cases}$	$F(x) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases}$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$
$G(p)$	$p_k = p(1-p)^k$	$1-(1-p)^{k+1}$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
$Esp(\lambda)$	$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$	$F(x) = \begin{cases} 1-\mathrm{e}^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda}^2$
$N(\mu,\sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$		μ	σ^2
$\Gamma(\alpha,\lambda)$	$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\mathrm{e}^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$		$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$