Esercizi di Calcolo delle Probabilità della 7^a settimana (Corso di Laurea in Matematica, Università degli Studi di Padova).

Esercizio 1. Sia $(X_n)_{n\in\mathbb{N}}$ un processo stocastico definito da $X_0 := x \in \mathbb{R}$ e, per ogni $n \in \mathbb{N}$,

$$X_{n+1} := aX_n + W_{n+1}$$

con $a \in \mathbb{R}$ e $(W_n)_n$ i.i.d. di legge $N(0, \sigma^2)$.

1. Dimostrare che, per ogni $n \in \mathbb{N}$, X_n è indipendente da W_{n+1}.
2. Dimostrare che $X_n \sim N(a^n x, \sigma^2(1 + a^2 + \ldots + a^{2n-2})$.
3. Per quali $a \in \mathbb{R}$ la successione $(X_n)_n$ converge in legge? A quale limite?

Esercizio 2. Siano $(X_n)_n$ variabili aleatorie tali che Var $[X_n] < c$ per ogni $n \in \mathbb{N}$ e per qualche $c \in \mathbb{R}^+$. Mostrare che $(X_n)_n$ soddisfa la legge debole dei grandi numeri se il **coefficiente di correlazione**, definito da

$$\rho(X_i, X_j) := \frac{\text{Cov }(X_i, X_j)}{\sqrt{\text{Var }[X_i]\text{Var }[X_j]}}$$

soddisfa una delle due relazioni

1. $\rho(X_i, X_j) \leq 0$ per ogni $i \neq j$;
2. $\rho(X_i, X_j) \to 0$ per $|i - j| \to \infty$

(suggerimento: ricordarsi come si dimostra la legge debole per variabili aleatorie scorrelate).

Esercizio 3. L'intervallo $[0, 1]$ è partizionato in n sottointervalli disgiunti $(I_i)_{i=1,\ldots,n}$, ciascuno di lunghezza p_i. Definiamo **entropia** di questa partizione la quantità

$$h := -\sum_{i=1}^{n} p_i \log p_i$$

Siano poi $(X_n)_n$ i.i.d. di legge $U(0, 1)$, e definiamo $Z_m(i)$ come il numero delle $X_1, \ldots, X_m$ che cadono nell'intervallo i-esimo della partizione. Definiamo infine

$$R_m := \prod_{i=1}^{n} p_i^{Z_m(i)}$$

1. Dare una definizione matematica delle $Z_m(i)$, $i = 1, \ldots, n$, $m \geq 1$.
2. Dimostrare che $\frac{\log R_m}{m} \to -h$ quasi certamente per $m \to \infty$.

Esercizio 4. Siano $U^n = (U_1, \ldots, U_n)$, $V^n = (V_1, \ldots, V_n)$ le coordinate di due punti scelti all'interno del cubo unitario n-dimensionale in modo indipendente e con legge uniforme, e sia X_n la distanza euclidea tra questi due punti.

1. Dimostrare che $\mathbb{E}[(U_i - V_i)^2] = \frac{1}{6}$ per ogni $i = 1, \ldots, n$.
2. Dimostrare che $X_n/\sqrt{n} \xrightarrow{\mathbb{P}} \frac{1}{\sqrt{6}}$ per $n \to \infty$.
3. Dimostrare che $\mathbb{E}[X_n]/\sqrt{n} \to \frac{1}{\sqrt{6}}$ per $n \to \infty$.

Esercizio 5. Il tempo di funzionamento (in giorni) di un componente prima di guastarsi è una variabile aleatoria di densità $f(x) := 2x\mathbf{1}_{(0,1)}(x)$ rispetto alla misura di Lebesgue. Supponiamo che appena questi componenti si guastano vengano rimpiazzati, e denotiamo con X_i il tempo di vita dell'i-esimo componente, e con $S_n := \sum_{i=1}^n X_i$ il momento dell'n-esimo guasto.

1. Calcolare $\mathbb{E}[X_i]$.
2. Supponendo che le $(X_i)_i$ siano indipendenti, calcolare il tasso di guasto a lungo termine
$$r := \lim_{n\to\infty} \frac{n}{S_n}$$
3. Quanti componenti sono necessari perchè la scorta, con probabilità pari al 90%, sia sufficiente per almeno 35 giorni?

Soluzioni su `http://www.math.unipd.it/~vargiolu/CalPro/`

Soluzioni

Esercizio 1.

1. Per induzione: X_0 è costante, quindi $\sigma(X_0) = \{\emptyset, \Omega\}$ ed X_0 è indipendente da W_1. Se $n \geq 1$ si ha che X_n dipende solo da $W_1, \ldots, W_n$, cioè $\sigma(X_n) \subseteq \sigma(W_1, \ldots, W_n)$, e questo implica che X_{n+1} dipende da X_n e W_{n+1}, quindi $\sigma(X_{n+1}) \subseteq \sigma(X_n, W_{n+1}) \subseteq (W_1, \ldots, W_n)$: questo significa che, per ogni $n \geq 1$, X_n è indipendente da W_{n+1}.

2. Calcolando le funzioni caratteristiche delle X_n si ha che $\varphi_n := \varphi_{X_n}$ è definita in modo ricorsivo in questo modo: $\varphi_0(t) = e^{itx}$, e per ogni $n \geq 1$,
$$\varphi_{n+1}(t) = \varphi_{aX_n}(t)\varphi_{W_{n+1}}(t) = \varphi_n(at)e^{-\frac{1}{2}\sigma^2 t^2}$$
ed è facile verificare per induzione che $\varphi_n(t) = \exp(ia^n xt - \frac{1}{2}\sigma^2(1 + a^2 + \ldots + a^{2n-2})t^2)$, che implica la tesi.

3. Usando il teorema di Paul Levy, vediamo che se $|a| \geq 1$ allora $\varphi_n \to \mathbf{1}_{\{0\}}(t)$ puntualmente, che non è una funzione caratteristica poiché è discontinua, quindi in questo caso $(X_n)_n$ non converge in legge. Viceversa, se $|a| < 1$, si ha che per ogni $t \in \mathbb{R}$
$$\varphi_n(t) = \exp\left(ia^n xt - \frac{1}{2}\sigma^2 \frac{1 - a^{2n}}{1 - a^2}\right) \to \exp\left(-\frac{1}{2}\sigma^2 \frac{1}{1 - a^2}\right)$$
e quindi, se $|a| < 1$, si ha che $X_n \rightharpoonup N(0; \frac{\sigma^2}{1-a^2})$.

Esercizio 2. Usando la diseguaglianza di Chebicev, abbiamo che per ogni $\delta > 0$
$$\mathbb{P}\left\{\left|\frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{n}\right| > \delta\right\} = \mathbb{P}\left\{\left|\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right| > \delta n\right\} \leq$$
$$\leq \quad \frac{\text{Var }[\sum_{i=1}^n X_i]}{\delta^2 n^2} = \frac{\sum_{i=1}^n \text{Var }[X_i] + 2\sum_{i<j} \text{Cov }(X_i, X_j)}{\delta^2 n^2} \leq \frac{c}{\delta^2 n} + \frac{2\sum_{i<j} \text{Cov }(X_i, X_j)}{\delta^2 n^2}$$

1. Se $\rho(X_i, X_j) < 0$, allora anche Cov $(X_i, X_j) < 0$, e si ha:
$$\mathbb{P}\left\{\left|\frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{n}\right| > \delta\right\} \leq \frac{c}{\delta^2 n} \to 0$$

2. L'ipotesi significa che per ogni $\varepsilon > 0$ esiste $I \in \mathbb{N}$ tale che $|i-j| > I$ implica $\rho(X_i, X_j) < \varepsilon$, e quindi Cov $(X_i, X_j) \leq \varepsilon c$. Ricordiamo che in ogni caso $\rho(X_i, X_j) < 1$, quindi Cov $(X_i, X_j) \leq c$. Allora
$$\mathbb{P}\left\{\left|\frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{n}\right| > \delta\right\} \leq$$
$$\leq \quad \frac{c}{\delta^2 n} + 2\frac{\sum_{i<j,|i-j|>I} \text{Cov }(X_i, X_j)}{\delta^2 n^2} + 2\frac{\sum_{i<j,|i-j|\leq I} \text{Cov }(X_i, X_j)}{\delta^2 n^2} \leq$$
$$\leq \quad \frac{c}{\delta^2 n} + \frac{n(n-1)\varepsilon c}{\delta^2 n^2} + 2\frac{nIc}{\delta^2 n^2} \to \frac{\varepsilon c}{\delta^2}$$
Siccome ε può essere preso piccolo a piacere, si ha la tesi.

Esercizio 3.

1. Si può formalizzare matematicamente la definizione che abbiamo dato in questo modo:

$$Z_m(i) := \sum_{j=1}^{m} \mathbf{1}_{\{X_j \in I_i\}}$$

In questo modo $Z_m(i)$ è uguale al numero di X_j, $j = 1, \ldots, m$ che cadono nell'intervallo I_i.

2. Abbiamo che

$$\frac{\log R_m}{m} = \frac{\sum_{i=1}^{n} Z_m(i) \log p_i}{m} = \sum_{i=1}^{n} \log p_i \frac{\sum_{j=1}^{m} \mathbf{1}_{X_j \in I_i}}{m}$$

Abbiamo poi che per ogni i, le $(\mathbf{1}_{X_j \in I_i})_j$ sono i.i.d. di legge $Be(p_i)$, poichè $\mathbb{P}\{\mathbf{1}_{X_j \in I_i} = 1\} = \mathbb{P}\{X_j \in I_i\} = \lambda(I_i) = p_i$, dove $\lambda(I_i)$ è la misura di Lebesgue dell'intervallo I_i. Per la legge forte dei grandi numeri di Kolmogorov-Khintchine, abbiamo che $\frac{Z_m(i)}{m} \overset{\text{q.c.}}{\to} p_i$, e siccome la convergenza quasi certa commuta con la somma e con la moltiplicazione per scalare, si ha che

$$\frac{\log R_m}{m} = \sum_{i=1}^{n} \log p_i \frac{Z_m(i)}{m} \quad \overset{\text{q.c.}}{\longrightarrow} \quad \sum_{i=1}^{n} p_i \log p_i = -h$$

Esercizio 4.

1. Siccome U_i e V_i sono indipendenti e di legge $U(0,1)$, abbiamo

$$\mathbb{E}[(U_i - V_i)^2] = \int_{[0,1]^2} (u-v)^2 \; du \; dv = \frac{1}{6}$$

o anche

$$\mathbb{E}[(U_i - V_i)^2] = \mathbb{E}[U_i^2] + \mathbb{E}[V_i^2] - 2\mathbb{E}[U_i]\mathbb{E}[V_i] = \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{6}$$

2. Abbiamo che

$$\frac{X_n^2}{n} = \frac{1}{n} \sum_{i=1}^{n} (U_i - V_i)^2$$

è la media delle variabili aleatorie i.i.d. $(U_i - V_i)^2$, che per il punto 1. hanno media uguale a $\frac{1}{6}$. Pertanto, per la legge forte dei grandi numeri si ha che $X_n^2/n \overset{\text{q.c.}}{\to} \frac{1}{6}$. Applicando la radice quadrata ad entrambi i membri, che è una funzione continua, abbiamo $X_n/\sqrt{n} \overset{\text{q.c.}}{\to} \frac{1}{\sqrt{6}}$, che implica la tesi.

3. Usiamo il teorema di convergenza dominata: abbiamo che $|U_i - V_i| \leq 1$ per ogni $i \geq 1$, quindi $X_n^2 = \sum_{i=1}^{n} (U_i - V_i)^2 \leq n$, e quindi $X_n/\sqrt{n} \leq 1$ per ogni $n \geq 1$, che appartiene ad L^1. Siccome dal punto 2. abbiamo $X_n/\sqrt{n} \overset{\text{q.c.}}{\to} \frac{1}{\sqrt{6}}$, segue la tesi.

Esercizio 5.

1. Dato che le X_i sono variabili aleatorie assolutamente continue e limitate, la loro speranza si calcola come
$$\mathbb{E}[X_i] = \int_0^1 x \cdot 2x\, dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$$

2. Siccome le $(X_i)_i$ sono i.i.d. e in L^1, per la legge forte dei grandi numeri di Kolmogorov-Khintchine abbiamo che $S_n/n \to \mathbb{E}[X_i] = \frac{2}{3}$ quasi certamente. Per le proprietà della convergenza quasi certa, abbiamo che
$$r = \lim_{n\to\infty} \frac{n}{S_n} = \frac{1}{\lim\limits_{n\to\infty} \frac{S_n}{n}} = \frac{1}{\frac{2}{3}} = \frac{3}{2}$$

3. Innanzitutto calcoliamo
$$\mathbb{E}[X_i^2] = \int_0^1 x^2 \cdot 2x\, dx = \left[\frac{2}{4}x^4\right]_0^1 = \frac{1}{2}$$
e quindi Var $[X_i] = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \frac{1}{2} - (\frac{2}{3})^2 = \frac{1}{18}$. Dato che sicuramente $X_i \leq 1$, abbiamo che sicuramente $S_n \geq 35$ implica $n \geq 35$, e quindi n dovrebbe essere abbastanza grande da poter applicare il teorema limite centrale. Abbiamo allora
$$0.90 = \mathbb{P}\{S_n \geq 35\} = \mathbb{P}\left\{\frac{S_n - n\mathbb{E}[X_i]}{\sqrt{n\text{Var }[X_i]}} \geq \frac{35 - \frac{2}{3}n}{\sqrt{\frac{1}{18}n}}\right\} \simeq 1 - \Phi\left(\frac{35 - \frac{2}{3}n}{\sqrt{\frac{1}{18}n}}\right)$$
dove Φ è la funzione di ripartizione di una legge $N(0,1)$. Da questo segue
$$\Phi\left(\frac{35 - \frac{2}{3}n}{\sqrt{\frac{1}{18}n}}\right) \simeq 0.1$$
cioè
$$\frac{35 - \frac{2}{3}n}{\sqrt{\frac{1}{18}n}} \simeq q_{0.1} = -q_{0.9} = -1.29$$
e quindi
$$35 \cdot 3 - 2n = -1.29\sqrt{\frac{1}{2}n} = -\frac{1.29}{2}\sqrt{2n} = -0.909\sqrt{n}$$
Risolvendo l'equazione di secondo grado in $\sqrt{n}$, si trovano le soluzioni $\sqrt{n} = \frac{1.075 \pm 29}{4}$. Siccome chiaramente bisogna prendere la soluzione positiva, si ha $\sqrt{n} \simeq 7.5$ e $n \simeq 7.5^2 = 56.25$. Ne segue che per $n \geq 57$ si ha la condizione richiesta.