CdL in Matematica, Probabilità (070AA)
A.A. 2021/22 - Prova scritta (unica prova in itinere) 2021-12-14

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Siano $\{a_n\}_n$, $\{b_n\}_n$ due successioni di numeri strettamente positivi e per ogni n siano X_n, Y_n variabili aleatorie indipendenti con densità esponenziale rispettivamente di parametri a_n, b_n. Posta $Z_n = X_n + Y_n$, dimostrare le seguenti affermazioni.

1. La famiglia delle leggi $\{P_{Z_n}\}_n$ è tesa se e solo se $\inf_n \min\{a_n, b_n\} > 0$.
2. Se entrambe le successioni $\{Z_n\}_n$, $\{\mathbb{E}[X_n|Z_n]\}_n$ convergono in legge allora entrambi i limiti $\lim_{n\to\infty} a_n$, $\lim_{n\to\infty} b_n$ esistono in $(0, +\infty]$ (ossia sono ciascuno un reale strettamente positivo oppure infinito).
3. Se $a_n = b_n$ per ogni n e $\lim_{n\to\infty} a_n$ esiste in $(0, \infty]$, allora per ogni $h \in C_b(\mathbb{R})$ la successione $\{\mathbb{E}[h(X_n)|Z_n]\}_n$ converge in legge. *(Sugg: trovare una espressione "esplicita" per* $\mathbb{E}[h(X_n)|Z_n]$ *come funzione di* Z_n*)*

Una soluzione:

1. Se $\inf_n \min\{a_n, b_n\} = \delta > 0$ allora si ha per ogni n

$$\mathbb{E}[X_n] = \frac{1}{a_n} \leq \frac{1}{\delta}, \quad \mathbb{E}[Y_n] = \frac{1}{b_n} \leq \frac{1}{\delta}$$

e di conseguenza

$$\mathbb{E}[|Z_n|] = \mathbb{E}[Z_n] = \mathbb{E}[X_n + Y_n] \leq \frac{2}{\delta}.$$

Segue dalla diseguaglianza di Markov che le leggi delle Z_n sono una famiglia tesa: per ogni $\varepsilon > 0$,

$$\sup_n P(|Z_n| > M) \leq \frac{2}{\delta M} \leq \varepsilon$$

se $M = \frac{2}{\delta\varepsilon}$.

Viceversa, se le leggi delle Z_n sono una famiglia tesa, per ogni ε esiste $M > 0$ tale che

$$\sup_n P(|Z_n| > M) \leq \varepsilon.$$

Poiché $0 \leq X_n \leq Z_n$ (la variabile Y_n è non negativa q.c.), si ha, per ogni n,

$$P(|X_n| > M) \leq P(|Z_n| > M) \leq \varepsilon.$$

Di conseguenza la famiglia delle leggi delle X_n è tesa. Ne segue che $\inf_n a_n > 0$. Precisamente, se esistesse una successione $a_{n_k} \to 0$ allora si avrebbe, per $k \to \infty$,

$$P(X_{n_k} \geq M) = e^{-a_{n_k} M} \to 1$$

in contraddizione con quanto mostrato sopra, se $\varepsilon < 1$. Ragionando allo stesso modo con le $\{Y_n\}_n$ segue che anche $\inf_n b_n > 0$.

3. Supponiamo che le $\{Z_n\}_n$ e le $\{\mathbb{E}[X_n|Z_n]\}_n$ convergano in legge. In particolare dal punto 1, poiché le leggi P_{Z_n} sono una famiglia tesa, si ha $\inf_n \min\{a_n, b_n\} = \delta > 0$. Inoltre, ricordando che la varianza di una variabile esponenziale di parametro λ è $1/\lambda^2$, troviamo che

$$\mathrm{Var}\, Z_n = \mathrm{Var}\, X_n + \mathrm{Var}\, Y_n = \frac{1}{a_n^2} + \frac{1}{b_n^2}$$

e quindi

$$\mathbb{E}\left[Z_n^2\right] = \mathrm{Var}\, Z_n + \mathbb{E}\left[Z_n\right]^2 = \frac{1}{a_n^2} + \frac{1}{b_n^2} + \left(\frac{1}{a_n} + \frac{1}{b_n}\right)^2 \leq \frac{6}{\delta^2}.$$

È noto dagli esercizi svolti nel corso che se una successione di variabili aleatorie converge in legge ed è limitata in L^p per qualche $p > 1$ allora anche i valori attesi passano al limite (che esiste finito). Troviamo quindi che il limite

$$\lim_n \mathbb{E}\left[Z_n\right] = \lim_n \left(a_n^{-1} + b_n^{-1}\right)$$

esiste finito (eventualmente nullo). D'altra parte si ha che $0 \leq X_n \leq Z_n$ quindi

$$0 \leq \mathbb{E}\left[X_n|Z_n\right] \leq Z_n, \quad \text{da cui} \quad \mathbb{E}\left[\left(\mathbb{E}\left[X_n|Z_n\right]\right)^2\right] \leq \mathbb{E}\left[Z_n^2\right]$$

e quindi anche $\mathbb{E}\left[X_n|Z_n\right]$ è uniformemente limitata in L^2. Passando al valore atteso, il simite

$$\lim_n \mathbb{E}\left[\mathbb{E}\left[X_n|Z_n\right]\right] = \lim_n \mathbb{E}\left[X_n\right] = \lim_n a_n^{-1}$$

esiste finito (eventualmente nullo). Per differenza anche $\lim_n b_n^{-1}$ esiste finito eventualmente nullo. Segue quindi che i limiti $\lim_n a_n$ e $\lim_n b_n$ esistono in $(0, \infty]$.

4. Iniziamo notando che le Z_n convergono in legge perché ad esempio, scrivendone la funzione caratteristica, troviamo

$$\varphi_{Z_n}(t) = \varphi_{X_n}(t)\varphi_{Y_n}(t) = \frac{a_n}{a_n + it}\frac{a_n}{a_n + it} \to \left(\frac{a}{a + it}\right)^2$$

intepretando il quoziente $a/(a + it) = 1$ se $a = \infty$. Il teorema di Paul Levy garantisce quindi la convergenza in legge (il limite puntuale è continuo in $t = 0$ anche nel caso $a = \infty$). Data $h \in C_b(\mathbb{R})$ basta mostrare che vale l'identità

$$\mathbb{E}\left[h(X_n)|Z_n\right] = \frac{1}{Z_n}\int_0^{Z_n} h(x)dx = g(Z_n).$$

dove poniamo

$$g(z) = \begin{cases} \frac{1}{z}\int_0^z h(x)dx & \text{se } z > 0, \\ h(0) & \text{se } z \leq 0. \end{cases}$$

Essendo h continua anche g lo è: la continuità è ovvia per $z \neq 0$, in $z = 0$ si verifica usando il teorema del valor medio integrale. Avendo mostrato che Z_n converge in legge (diciamo verso una variabile Z), segue da fatti visti nel corso che anche $g(Z_n)$

converge in legge verso $g(Z)$ (basta osservare che per ogni $f \in C_b(\mathbb{R})$ la composizione $f(g(Z_n)) = f \circ g(Z_n)$ e $f \circ g \in C_b(\mathbb{R})$).

Per mostrare l'identità sopra, evitiamo di scrivere l'indice n per semplicità e supponiamo quindi X, Y esponenziali di parametro a indipendenti e $Z = X + Y$. Allora per ogni φ boreliana e limitata,

$$\begin{aligned}\mathbb{E}\left[h(X)\varphi(Z)\right] &= \int_0^\infty \int_0^\infty h(x)\varphi(x+y)e^{-ax-ay}a^2 dxdy \\ &= \int_0^\infty \int_x^\infty h(x)\varphi(z)e^{-ax-a(z-x)}a^2 dxdz \\ &= \int_0^\infty \varphi(z)e^{-az}a^2 \int_0^z h(x)dxdz.\end{aligned}$$

Ponendo $h(x) = 1$ costante, si trova

$$\mathbb{E}\left[\varphi(Z)\right] = \int_0^\infty \varphi(z)e^{-az}a^2 zdz,$$

ossia che la densità di Z è data dalla funzione

$$z \in [0,\infty) \mapsto a^2 e^{-az} z.$$

(è una densità $\Gamma(2,a)$). Di conseguenza troviamo che

$$\mathbb{E}\left[h(X)\varphi(Z)\right] = \int_0^\infty \left[\frac{1}{z}\int_0^z h(x)dx\right]\varphi(z)a^2 e^{-az}zdz.$$

Essendo φ arbitraria questo mostra l'identità cercata.

Problema 2

Siano $(X_n)_{n=1}^\infty$ variabili aleatorie reali i.i.d. gaussiane standard $\mathcal{N}(0,1)$. Si definisca per ogni $n \geq 1$ la variabile

$$M_n = \max_{i=1,\ldots,n} X_i.$$

Usando le diseguaglianze (non serve dimostrarle)

$$\frac{e^{-z^2/2}}{\sqrt{2\pi}}\left(\frac{1}{z} - \frac{1}{z^3}\right) \leq P(Z > z) \leq \frac{e^{-z^2/2}}{z\sqrt{2\pi}}$$

valide per ogni $z > 0$ e Z variabile gaussiana reale standard, mostrare che

1. P-q.c. vale $\limsup_{n\to\infty} \frac{X_n}{\sqrt{\log n}} = \sqrt{2}$,
2. vale $\lim_{n\to\infty} M_n/\sqrt{\log n} = \sqrt{2}$ in probabilità *(Sugg: considerare la funzione di ripartizione di M_n)*,
3. il limite in probabilità $\lim_{n\to\infty}(M_n - X_n)/\sqrt{\log n}$ esiste e calcolarlo.

Una soluzione:

1. Per $\varepsilon > 0$ consideriamo l'evento

$$\left\{ \frac{X_n}{\sqrt{\log n}} > \sqrt{2+\varepsilon} \right\}$$

che ha probabilità (usando la disuguaglianza dall'alto nel suggerimento)

$$P\left(\frac{X_n}{\sqrt{\log n}} > \sqrt{2+\varepsilon} \right) \leq \frac{e^{-\log n \frac{2+\varepsilon}{2}}}{\sqrt{(2+\varepsilon)\log n}\sqrt{2\pi}} = \frac{1}{n^{1+\varepsilon/2}\sqrt{\log n}\sqrt{(2+\varepsilon)2\pi}}.$$

Essendo tale probabilità sommabile, il lemma di Borel-Cantelli (prima parte) garantisce che

$$P\left(\limsup_n \left\{ \frac{X_n}{\sqrt{\log n}} > \sqrt{2+\varepsilon} \right\} \right) = 0,$$

ossia che P-q.c., si ha definitivamente in n

$$\frac{X_n}{\sqrt{\log n}} \leq \sqrt{2+\varepsilon},$$

ossia $\limsup_n \frac{X_n}{\sqrt{\log n}} \leq \sqrt{2+\varepsilon}$. Ripetendo il ragionamento ma con gli eventi (indipendenti)

$$\left\{ \frac{X_n}{\sqrt{\log n}} > \sqrt{2-\varepsilon} \right\},$$

si trova che le loro probabilità non sono sommabili, usando la diseguaglianza dal basso nel suggerimento:

$$P\left(\frac{X_n}{\sqrt{\log n}} > \sqrt{2-\varepsilon} \right) \geq \frac{1}{n^{1-\varepsilon/2}\sqrt{2\pi}} \left(\frac{1}{\sqrt{(2-\varepsilon)\log n}} - \frac{1}{((2-\varepsilon)\log n)^{3/2}} \right)$$

e di conseguenza la seconda parte del lemma di Borel-Cantelli garantisce che P-q.c.

$$\limsup_n \frac{X_n}{\sqrt{\log n}} \geq \sqrt{2-\varepsilon}.$$

2. Osserviamo innanzitutto che essendo $X_n \leq M_n$, si ha dal punto precedente che P-q.c.

$$\liminf_n \frac{M_n}{\sqrt{\log n}} \geq \sqrt{2}.$$

Per ottenere la convergenza in probabilità è quindi sufficiente mostrare che per ogni $\varepsilon > 0$, si abbia

$$\lim_{n\to\infty} P\left(\frac{M_n}{\sqrt{\log n}} > \sqrt{2} + \varepsilon \right) = 0,$$

ossia che

$$\lim_{n\to\infty} P\left(\frac{M_n}{\sqrt{\log n}} \leq \sqrt{2} + \varepsilon \right) = 1.$$

Per ogni $x \geq 0$, calcoliamo come da suggerimento la funzione di ripartizione di M_n in x,

$$\begin{aligned} P(M_n \leq x) &= P(\max_{i=1,\dots,n} X_i \leq x) = P(X_i \leq x \quad \forall i = 1,\dots,n) \\ &= \prod_{i=1}^{n} P(X_i \leq x) = (1 - P(X_1 > x))^n. \end{aligned}$$

Posto $x = (\sqrt{2}+\varepsilon)\sqrt{\log n}$, per $\varepsilon > 0$, troviamo che

$$\begin{aligned} P(M_n/\sqrt{\log n} \leq \sqrt{2}+\varepsilon) &= \Big(1 - P(X_1 > (\sqrt{2}+\varepsilon)\log n)\Big)^n \\ &\geq \left(1 - o\left(\frac{1}{n}\right)\right)^n \to 1 \end{aligned}$$

per $n \to \infty$, avendo usato la limitazione dall'alto

$$P(X_1 > (\sqrt{2}+\varepsilon)\log n) \leq \frac{1}{n^{(\sqrt{2}+\varepsilon)^2/2}\sqrt{2\pi}} \frac{1}{(\sqrt{2}+\varepsilon)\sqrt{\log n}} = o\left(\frac{1}{n}\right).$$

3. Usando il punto precedente, basta notare che le variabili $X_n/\sqrt{\log n}$ convergono verso zero in probabilità, perché per ogni $\varepsilon > 0$

$$P(|X_n|/\sqrt{\log n} > \varepsilon) = P(|X_1| > \varepsilon\sqrt{\log n}) \to 0.$$

Dato che la somma (in questo caso la differenza) di variabili che convergono in probabilità converge pure in probabilità, otteniamo che

$$\lim_{n\to\infty} \frac{M_n - X_n}{\sqrt{\log n}} = \sqrt{2}$$

in probabilità.

CdL in Matematica, Probabilità (070AA)
A.A. 2021/22 - Prova scritta 2022-01-26

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Sia $(X_i)_{i=1}^{\infty}$ una successione di variabili aleatorie reali indipendenti e identicamente distribuite, assolutamente continue, con densità

$$f(x) = \begin{cases} \frac{c}{x^2 \log|x|} & \text{se } |x| \geq 2, \\ 0 & \text{altrimenti,} \end{cases}$$

dove $c > 0$ è una opportuna costante. Si definiscano inoltre, per ogni $n \geq 1$, le variabili

$$S_n = \sum_{i=1}^{n} X_i \quad \text{ed} \quad S_n^* = \sum_{i=1}^{n} X_i I_{\{|X_i|<n\}}.$$

1. Argomentare che $c > 0$ esiste finita (non è richiesto di calcolarla) e dire per quali $k \in \mathbb{R}$ si ha $\mathbb{E}\left[|X_1|^k\right] < \infty$.
2. Mostrare il limite in probabilità $\lim_{n\to\infty} S_n^*/n = 0$ *(Sugg: considerare la varianza di S_n^*)*
3. Mostrare che $\lim_{n\to\infty} P\left(S_n \neq S_n^*\right) = 0$ e dedurne il limite in probabilità
$$\lim_{n\to\infty} \frac{S_n}{n} = 0.$$
4. *(Facoltativo)* L'ultimo limite mostrato al punto sopra vale anche quasi certamente?

Una soluzione:

1. Si ha
$$\mathbb{E}\left[|X_1|^k\right] = 2c\int_2^{\infty} x^k \cdot \frac{1}{x^2 \log x} dx = 2c \int_2^{\infty} \frac{1}{x^{1+(1-k)} \log x} dx.$$
Se $k < 1$ l'integrale è convergente per confronto con l'integrale di $1/x^{1+\varepsilon}$ per un qualsiasi $0 < \varepsilon < 1-k$. Se $k = 1$ si ottiene l'integrale di $1/(x \log x) = (\log\log x)'$ che diverge. Lo stesso accade per qualsiasi $k > 1$ (altrimenti ne seguirebbe l'integrabilità di X_1).

2. Seguendo il suggerimento, osserviamo che S_n^* è la somma delle n variabili indipendenti e identicamente distribuite $\left\{X_i I_{\{|X_i|<n\}}\right\}_{i=1}^n$. Dato che la densità di X_i è simmetrica la legge di ciascuna X_i è uguale a quella di $-X_i$, otteniamo l'identità

$$\mathbb{E}\left[X_i I_{\{|X_i|<n\}}\right] = \mathbb{E}\left[-X_i I_{\{|-X_i|<n\}}\right] = -\mathbb{E}\left[X_i I_{\{|X_i<n|\}}\right] = 0,$$

dove abbiamo usato anche il fatto che $X_i I_{\{|X_i|<n\}}$ è integrabile (essendo limitata). Calcoliamo poi

$$\operatorname{Var}\left(X_i I_{\{|X_i|<n\}}\right) = \mathbb{E}\left[X_i^2 I_{\{|X_i<n|\}}\right] = 2c\int_2^n \frac{1}{\log x} dx.$$

Usando l'indipendenza, segue che

$$\operatorname{Var}(S_n^*) = \frac{1}{n^2}\sum_{i=1}^n \operatorname{Var}\left(X_i I_{\{|X_i|<n\}}\right) = \frac{2c\int_2^n \frac{1}{\log x}dx}{n}$$

L'ultima quantità è infinitesima per $n \to \infty$, ad esempio per il teorema di Cesaro, essendo i termini $\int_k^{k+1} 1/\log x dx$ infinitesimi, oppure usando il teorema di de l'Hôpital. Segue pertanto che $S_n^* \to 0$ in L^2 e quindi in probabilità.

3. Stimiamo

$$\begin{aligned} P(S_n^* \neq S_n) &\leq P\left(\exists i \in \{1,\dots,n\} \,:\, X_i \neq X_i I_{\{|X_i|<n\}}\right) \\ &\leq \sum_{i=1}^n P(|X_i| \geq n) \\ &= nP(|X_1| \geq n) \\ &= 2cn\int_n^\infty \frac{1}{x^2\log x}dx. \end{aligned}$$

Anche in questo caso possiamo usare il teorema di de l'Hôpital:

$$\lim_{n\to\infty} \frac{\int_n^\infty \frac{1}{x^2\log x}dx}{1/n} = \lim_{n\to\infty} \frac{-\frac{1}{n^2\log n}}{-\frac{1}{n^2}} = 0.$$

In particolare segue che per ogni $\varepsilon > 0$,

$$P(|S_n^* - S_n| > \varepsilon) \leq P(S_n^* \neq S_n) \to 0,$$

quindi $\lim_{n\to\infty}(S_n^* - S_n) = 0$ in probabilità, e a maggior ragione $\lim_{n\to\infty}(S_n^* - S_n)/n = 0$. Combinando con il punto precedente si trova il limite in probabilità

$$\lim_{n\to\infty}\frac{S_n}{n} = \lim_{n\to\infty}\frac{(S_n - S_n^*)}{n} + \lim_{n\to\infty}\frac{S_n^*}{n} = 0.$$

Problema 2

Siano X, Y variabili aleatorie reali e sia $f : \mathbb{R} \to \mathbb{R}$ boreliana.

1. Mostrare che X è indipendente da Y se e solo se, per ogni $t \in \mathbb{R}$, $\mathbb{E}\left[e^{itX}|Y\right]$ è $\mathbb{P}$-q.c. costante. *(Sugg: considerare la funzione caratteristica della coppia* (X,Y)*)*
2. Mostrare che $X = f(Y)$ $\mathbb{P}$-q.c. se e solo se, per ogni $t \in \mathbb{R}$, vale $\mathbb{P}$-q.c.

$$\mathbb{E}\left[e^{itX}|Y\right] = e^{itf(Y)}.$$

3. Supponendo che (X,Y) sia un vettore gaussiano centrato con matrice delle covarianze

$$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix},$$

calcolare $\mathbb{E}\left[e^{itX}|Y\right]$, fornendo una espressione il più possibile esplicita come funzione di Y.

(Nota: le tre domande non sono necessariamente collegate tra loro.)

Una soluzione:

1. Se X è indipendente da Y, allora per risultati noti dalla teoria

$$\mathbb{E}\left[e^{itX}|Y\right] = \mathbb{E}\left[e^{itX}\right]$$

è costante $\mathbb{P}$-q.c. (per ogni $t \in \mathbb{R}$ fissato).

Viceversa, supponiamo che per ogni $t \in \mathbb{R}$ la variabile $\mathbb{E}\left[e^{itX}|Y\right]$ sia costante $\mathbb{P}$-q.c.: passando al valore atteso otteniamo che $\mathbb{P}$-q.c. si ha

$$\mathbb{E}\left[e^{itX}|Y\right] = \mathbb{E}\left[\mathbb{E}\left[e^{itX}|Y\right]\right] = \mathbb{E}\left[e^{itX}\right] = \varphi_X(t).$$

Consideriamo allora la funzione caratteristica della coppia di variabili (Y, X):

$$\begin{aligned}
\varphi_{(Y,X)}(s,t) &= \mathbb{E}\left[e^{i(sY+tX)}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[e^{i(sY+tX)}|Y\right]\right] = \mathbb{E}\left[e^{isY}\mathbb{E}\left[e^{itX}|Y\right]\right] \\
&= \mathbb{E}\left[e^{isY}\varphi_X(t)\right] \\
&= \varphi_Y(s)\varphi_X(t).
\end{aligned}$$

Per quanto visto nella teoria questa identità implica che X e Y siano indipendenti.

2. Se $X = f(Y)$, allora essendo X una variabile $\sigma(Y)$-misurabile segue che

$$\mathbb{E}\left[e^{itX}|Y\right] = e^{itX} = e^{itf(Y)}.$$

Per il viceversa, consideriamo la coppia di variabili (Y, X) e la sua funzione caratteristica:

$$\varphi_{(Y,X)}(s,t) = \mathbb{E}\left[e^{i(sY+tX)}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{i(sY+tX)}|Y\right]\right] = \mathbb{E}\left[e^{isY}\mathbb{E}\left[e^{itX}|Y\right]\right] = \mathbb{E}\left[e^{i(sY+tf(Y))}\right],$$

che riconosciamo essere anche la funzione caratteristica della coppia di variabili $(Y, f(Y))$:

$$\varphi_{(Y,f(Y))}(s,t) = \mathbb{E}\left[e^{i(sY+tf(Y))}\right] = \varphi_{(Y,X)}(s,t).$$

Ne segue che la legge congiunta di (Y, X) coincide con quella di $(Y, f(Y))$. Perciò

$$\begin{aligned}
P(|f(Y) - X| > 0) &= P((Y,X) \in \{(y,x) : |f(y) - x| > 0\}) \\
&= P((Y, f(Y)) \in \{(y,x) : |f(y) - x| > 0\}) \\
&= P(|f(Y) - f(Y)| > 0) = 0,
\end{aligned}$$

da cui $X = f(Y)$ $\mathbb{P}$-q.c.

3. Determiniamo prima $\alpha \in \mathbb{R}$ in modo che

$$\operatorname{Cov}(X - \alpha Y, Y) = 0,$$

e quindi che il vettore gaussiano $(X-\alpha Y, Y)$, trasformazione lineare del vettore (X, Y), abbia componenti non correlate e quindi indipendenti. Troviamo

$$\begin{aligned} 0 = \operatorname{Cov}(X-\alpha Y, Y) &= \operatorname{Cov}(X, Y) - \alpha \operatorname{Var} Y \\ &= 2 - 5\alpha \end{aligned}$$

da cui $\alpha = 2/5$. Sfruttiamo ora l'indipendenza per calcolare

$$\begin{aligned} \mathbb{E}\left[e^{itX}|Y\right] &= \mathbb{E}\left[e^{it(X-\alpha Y+\alpha Y)}|Y\right] \\ &= \mathbb{E}\left[e^{it(X-\alpha Y)}e^{it\alpha Y}|Y\right] \\ &= e^{it\alpha Y}\mathbb{E}\left[e^{it(X-\alpha Y)}\right] \\ &= e^{it2Y/5}e^{-\frac{t^2\sigma^2}{2}}, \end{aligned}$$

dove

$$\begin{aligned} \sigma^2 &= \operatorname{Var}(X-\alpha Y) = \operatorname{Var}(X) + \alpha^2 \operatorname{Var}(Y) - 2\alpha \operatorname{Cov}(X, Y) \\ &= 1 + \alpha^2 \cdot 5 - 4\alpha \\ &= 1 + \frac{4}{5} - \frac{8}{5} \\ &= \frac{1}{5}. \end{aligned}$$

Concludiamo che

$$\mathbb{E}\left[e^{itX}|Y\right] = e^{-it2Y/5 - \frac{t^2}{10}}.$$

CdL in Matematica, Probabilità (070AA)
A.A. 2021/22 - Prova scritta 2022-02-23

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Su uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ siano definite due variabili aleatorie reali X e Y con densità congiunta

$$f(x,y) = \begin{cases} c \cdot \frac{e^{-\frac{y^2}{2}}}{1+y} \cdot e^{-\frac{x}{1+y}} & \text{se } x > 0 \text{ e } y > 0, \\ 0 & \text{altrimenti,} \end{cases}$$

dove $c > 0$ è una opportuna costante.

1. Calcolare esplicitamente c.
2. Calcolare la covarianza tra le variabili Y e $X/(1+Y)$.
3. Esprimere come funzioni (esplicite) di Y le variabili aleatorie

$$\mathbb{E}[X|Y] \quad \text{e} \quad \operatorname{Var}(X|Y).$$

Una soluzione:

1. Deve valere $\int_{\mathbb{R}^2} f = 1$. Usando il teorema di Fubini per integrare prima rispetto alla variabile x, si ottiene

$$1 = \int_0^\infty c \cdot e^{-\frac{y^2}{2}} \int_0^\infty \frac{e^{-\frac{x}{1+y}}}{1+y} dx dy = \int_0^\infty c \cdot e^{-\frac{y^2}{2}} dy = c\frac{1}{2}\sqrt{2\pi},$$

avendo usato l'integrale gaussiano $\int_{-\infty}^{-\infty} e^{-y^2/2} dy = \sqrt{2\pi}$. Segue che $c = \sqrt{2/\pi}$.

2. Mostriamo che le variabili Y e $X/(1+Y)$ sono indipendenti (da questo segue che la covarianza è nulla, purché sia ben definita, ad esempio se le variabili ammettono momento secondo finito). È sufficiente mostrare che, per ogni s, $t \in \mathbb{R}$, si ha

$$\mathbb{E}\left[e^{i(sX/(1+Y)+tY)}\right] = \varphi(s)\psi(t),$$

per opportune funzioni $\varphi(s)$, $\psi(t)$ che poi riconosceremo come le funzioni caratteristiche delle due variabili. Usando la densità congiunta e il teorema di Fubini-Tonelli per integrare prima rispetto alla variabile X, troviamo

$$\mathbb{E}\left[e^{i(sX/(1+Y)+tY)}\right] = \int_0^\infty \sqrt{2/\pi} e^{-\frac{y^2}{2}} e^{ity} \int_0^\infty \frac{e^{-\frac{x}{1+y}}}{1+y} e^{isx/(1+y)} dx dy.$$

Nell'integrale rispetto alla variabile x effettuiamo il cambio di variabile $z = x/(1+y)$ ottenendo

$$\int_0^\infty \frac{e^{-\frac{x}{1+y}}}{1+y} e^{isx/(1+y)} dx = \int_0^\infty e^{-z} e^{isz} dz$$

che è la funzione caratteristica di una variabile con densità esponenziale di parametro 1 (abbiamo quindi identificato la legge di $X/(1+Y)$, che ha in particolare tutti i momenti finiti). Di conseguenza, introducendo la funzione caratteristica

$$\psi(t) = \int_0^\infty \sqrt{2/\pi} e^{-\frac{y^2}{2}} e^{ity} dy$$

segue che la densità di Y è $\sqrt{2/\pi} e^{-\frac{y^2}{2}} I_{\{y>0\}}$ (e quindi ammette tutti i momenti finiti) e l'indipendenza.

3. Sfruttiamo l'indipendenza osservata al punto precedente e le proprietà della speranza condizionale:

$$\begin{aligned}\mathbb{E}[X|Y] &= \mathbb{E}[(1+Y)\cdot X/(1+Y)|Y] = (1+Y)\mathbb{E}[X/(1+Y)|Y] \\ &= (1+Y)\mathbb{E}[\mathrm{Exp}(1)] = (1+Y).\end{aligned}$$

E similmente,

$$\begin{aligned}\mathbb{E}\left[X^2|Y\right] &= \mathbb{E}\left[(1+Y)^2\cdot X^2/(1+Y)^2|Y\right] = (1+Y)^2\mathbb{E}\left[X^2/(1+Y)^2|Y\right] \\ &= (1+Y)^2\mathbb{E}\left[(\mathrm{Exp}(1))^2\right] = (1+Y)^2 2.\end{aligned}$$

Da cui $\mathrm{Var}(X|Y) = (1+Y)^2$.

Problema 2

Su uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ sia definita $(X_n)_{n=1}^\infty$ una successione di variabili aleatorie reali i.i.d., dotate di momento secondo finito e centrate. Si ponga, per ogni $n \geq 1$,

$$Y_n = \sum_{1\leq i<j\leq n} X_i X_j.$$

Mostrare che, per $n \to \infty$,

i) $Y_n/n^2 \xrightarrow{L^2} 0$,

ii) $Y_n/n^2 \xrightarrow{\mathbb{P}-q.c.} 0$, *(Sugg: elevare al quadrato la variabile $S_n = \sum_{i=1}^n X_i$)*

iii) Y_n/n converge in legge (descrivere la legge del limite).

Una soluzione:

1. Poniamo $\sigma^2 = \mathbb{E}\left[X_1^2\right]$. Calcoliamo

$$\mathbb{E}\left[Y_n^2\right] = \mathbb{E}\left[\sum_{1\leq i<j\leq n}\sum_{1\leq k<\ell\leq n} X_i X_j X_k X_\ell\right].$$

Osserviamo che ciascuna variabile $X_i X_j X_k X_\ell$ è integrabile: bisogna distinguere alcuni casi: se i quattro indici i, j, k, ℓ sono tutti diversi, allora per indipendenza

$$\mathbb{E}[|X_i X_j X_k X_\ell|] = \mathbb{E}[|X_1|]^4 < \infty,$$

mentre se due indici coincidono e due indici sono diversi, allora

$$\mathbb{E}\left[|X_i X_j X_k X_\ell|\right] = \mathbb{E}\left[X_1^2\right]\mathbb{E}\left[|X_1|\right]^2 < \infty.$$

Se infine $i = k$ e $j = \ell$,

$$\mathbb{E}\left[|X_i X_j X_k X_\ell|\right] = \mathbb{E}\left[X_1^2\right]^2 < \infty.$$

Non può invece accadere che tre o più indici coincidano (viste le condizioni $i < j$ e $k < \ell$). Ripetendo la stessa casistica per i valori attesi (senza valore assoluto), segue che nel primo e nel secondo caso il valore atteso è nullo, quindi rimane solamente il terzo e

$$\mathbb{E}\left[Y_n^2\right] = \sum_{1 \le i < j \le n} \mathbb{E}\left[X_i^2 X_j^2\right] = \sum_{1 \le i < j \le n} \mathbb{E}\left[X_1^2\right]^2 = \binom{n}{2}\sigma^4.$$

Dividendo per n^4, si trova che

$$\mathbb{E}\left[(Y_n/n^2)^2\right] = \binom{n}{2}\sigma^4/n^4 \to 0$$

al tendere di $n \to \infty$, da cui la convergenza in L^2.

2. Per la convergenza quasi certa, partiamo dal suggerimento

$$(S_n)^2 = \sum_{i=1}^n X_i^2 + 2Y_n.$$

Sappiamo dalla legge forte dei grandi numeri che vale la convergenza quasi certa

$$\frac{S_n}{n} \to 0$$

e pure (usando la legge forte dei grandi numeri di Kolmogorov per variabili i.i.d. integrabili) quasi certamente

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \to \mathbb{E}\left[X_1^2\right] = \sigma^2,$$

e quindi

$$\frac{1}{n^2}\sum_{i=1}^n X_i^2 \to 0.$$

Dividendo per $2n^2$, troviamo che

$$\frac{Y_n}{n^2} = \frac{1}{2}\left(\frac{S_n}{n}\right)^2 - \frac{1}{2n^2}\sum_{i=1}^n X_i^2 \to 0.$$

3. Per il teorema limite centrale sappiamo che

$$\frac{S_n}{\sqrt{n}} \to Z\sigma$$

dove Z ha legge gaussiana standard $\mathcal{N}(0,1)$. Poiché elevare al quadrato è una funzione continua, vale la convergenza in legge

$$\left(\frac{S_n}{\sqrt{n}}\right)^2 \to Z^2\sigma^2.$$

Ragionando come nel punto precedente ma dividendo stavolta per n invece di n^2, troviamo che

$$\frac{Y_n}{n} = \frac{1}{2}\left(\frac{S_n}{\sqrt{n}}\right)^2 - \frac{1}{2n}\sum_{i=1}^{n} X_i^2 \to \frac{Z^2\sigma^2}{2} - \frac{\sigma^2}{2} = \frac{\sigma^2}{2}(Z^2-1),$$

dove abbiamo usato il fatto noto dal corso che se $U_n \to U$ in legge e $V_n \to c$ una costante in legge (che segue dalla convergenza quasi certa) allora $U_n + V_n \to U + c$.

CdL in Matematica, Probabilità (070AA)
A.A. 2021/22 - Prova scritta 2022-06-20

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Sia X una v.a. reale avente legge di Cauchy, ossia assolutamente continua con densità

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad \text{per } x \in \mathbb{R}.$$

1. Mostrare che la funzione caratteristica di X è $\varphi_X(t) = e^{-|t|}$.
2. Mostrare che $1/X$ è una variabile avente legge di Cauchy.
3. Sia Y una variabile aleatoria reale indipendente da X, con $|Y| \leq 1$ $\mathbb{P}$-q.c. e tale che Y/X abbia legge di Cauchy. Mostrare che $\mathbb{P}$-q.c. si ha $|Y| = 1$.
4. Siano Y, Z due variabili aleatorie reali tra loro indipendenti, con

$$P(Y = 1) = P(Y = -1) = 1/2$$

e tali che $1/Z$, Y/Z abbiano entrambe la stessa legge di Z. Mostrare che Z è simmetrica, ossia $-Z$ ha la stessa legge di Z. La legge di Z è necessariamente di Cauchy?

Una soluzione:

1. Scriviamo la formula di inversione:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-|t|}e^{-itx}dt = \frac{1}{\pi}\int_0^{\infty} e^{-t}\frac{e^{-itx}+e^{itx}}{2}dt$$
$$= \frac{1}{\pi}\int_0^{\infty} e^{-t}\cos(tx)dt.$$

Integrando per parti due volte si ottiene

$$\int_0^{\infty} e^{-t}\cos(tx)dt = \frac{1}{1+x^2}.$$

Per iniettività della funzione caratteristica, troviamo quindi che $e^{-|t|}$ è la funzione caratteristica di una variabile con legge di Cauchy.

2. Se X ha densità $f(x)$ e quindi funzione di ripartizione $F(x) = \int_0^x f(s)ds$, allora la funzione di ripartizione di $1/X$ è (per $t > 0$)

$$P(1/X < t) = P(X > 1/t) = 1 - F(1/t).$$

Derivando otteniamo per la densità l'espressione

$$\frac{1}{t^2}f(1/t).$$

Nel caso $f(x) = \frac{1}{\pi(1+x^2)}$, troviamo

$$\frac{1}{t^2} f(1/t) = \frac{1}{t^2 \pi (1+t^{-2})} = \frac{1}{\pi(1+t^2)}.$$

Per $t < 0$ si ragiona allo stesso modo (oppure si sfrutta che la densità è pari).

3. Calcoliamo la funzione caratteristica di $\frac{Y}{X}$. Dato che $\frac{1}{X}$ ha legge di Cauchy (ved. es. precedente), per indipendenza si ha

$$\phi_{\frac{Y}{X}}(t) = \mathbb{E}\left[e^{it\frac{Y}{X}}\right] = \int \mathbb{E}\left[e^{it\frac{y}{X}}\right] \mu_Y(dy) = \int e^{-|ty|} \mu_Y(dy) = \int_{|y|\leq 1} e^{-|ty|} \mu_Y(dy).$$

Per l'ipotesi abbiamo $\phi_{\frac{Y}{X}}(t) = \phi_X(t)$, e cioè

$$\int_{|y|\leq 1} e^{-|ty|} \mu_Y(dy) = e^{-|t|},$$

Scritta diversamente, questa relazione diventa

$$\int_{|y|\leq 1} \left(e^{-|ty|} - e^{-|t|}\right) \mu_Y(dy) = 0.$$

Dato che la funzione integranda è positiva, se ne deduce che μ_Y-q.o. y,

$$e^{-|ty|} = e^{-|t|},$$

da cui $|y| = 1$.

4. Sia ϕ_Z la funzione caratteristica di Z. Dato che $\frac{1}{Z}$ e $\frac{Y}{Z}$ hanno la stessa legge di X, la funzione caratteristica di $\frac{Y}{Z}$ è uguale a ϕ_Z ed inoltre

$$\phi_{\frac{Y}{Z}}(t) = \int \mathbb{E}\left[e^{i\frac{ty}{Z}}\right] \mu_Y(dy) = \int \phi_Z(ty) \mu_Y(dy) = \frac{1}{2}\phi_Z(t) + \frac{1}{2}\phi_Z(-t).$$

L'uguaglianza $\phi_Z(t) = \phi_{\frac{Y}{Z}}(t)$ può allora essere scritta nella forma

$$\phi_Z(t) = \frac{1}{2}\phi_Z(t) + \frac{1}{2}\phi_Z(-t),$$

quindi $\phi_Z(t) = \phi_Z(-t)$. D'altra parte $\phi_{-Z}(t) = \phi_Z(-t)$. Z non è necessariamente di Cauchy, ad esempio può essere $P(Z = 1) = P(Z = -1) = 1/2$ e le ipotesi sono comunque soddisfatte.

Problema 2

Sia $\alpha > 0$ un parametro reale e siano $(X_n)_{n=1}^{\infty}$ variabili aleatorie reali indipendenti, tali che per ogni n si abbia

$$P(X_n = n^\alpha - 1) = \frac{1}{n^\alpha}, \quad P(X_n = -1) = 1 - \frac{1}{n^\alpha}.$$

Si ponga $S_n = \sum_{k=1}^{n} X_k$. Dire se le seguenti affermazioni sono vere o false (giustificando opportunamente le risposte).

1. Se $\alpha < 1$, allora $\lim_{n\to\infty} S_n/n = 0$ in $L^2(\mathbb{P})$.
2. Se $\alpha < 1$, allora $\lim_{n\to\infty} S_n/n = 0$ $\mathbb{P}$-q.c.
3. Se $\alpha > 1$, allora $\lim_{n\to\infty} S_n/n = -1$ in $L^2(\mathbb{P})$.
4. Se $\alpha > 1$, allora $\lim_{n\to\infty} S_n/n = -1$ $\mathbb{P}$-q.c.
5. Se $\alpha = 1$, allora la serie $\sum_{n=1}^{\infty} X_n/n$ è convergente $\mathbb{P}$-q.c.

Una soluzione:

1. L'affermazione è vera. Infatti possiamo calcolare

$$\mathbb{E}[X_n] = (n^\alpha - 1)\cdot \frac{1}{n^\alpha} - 1 \cdot \left(1 - \frac{1}{n^\alpha}\right) = 0$$

e

$$\operatorname{Var}(X_n) = \mathbb{E}\left[X_n^2\right] = (n^\alpha - 1)^2 \cdot \frac{1}{n^\alpha} + 1^2 \cdot \left(1 - \frac{1}{n^\alpha}\right) = n^\alpha - 1.$$

Di conseguenza, per indipendenza

$$\mathbb{E}\left[S_n^2\right] = \sum_{k=1}^{n} \mathbb{E}\left[X_k^2\right] = \sum_{k=1}^{n} (k^\alpha - 1) = \sum_{k=1}^{n} k^\alpha - n.$$

Confrontando ad esempio con l'integrale $\int_1^n x^\alpha$, si trova che

$$\sum_{k=1}^{n} k^\alpha = O(n^{\alpha+1}),$$

da cui

$$\frac{\mathbb{E}\left[S_n^2\right]}{n^2} = O(n^{\alpha-1}) \to 0$$

per $n \to \infty$.

2. L'affermazione è vera per la legge forte dei grandi numeri (caso di variabili di quadrato integrabile) usando i calcoli del punto precedente, si trova che

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left[X_n^2\right]}{n^2} = \sum_{n=1}^{\infty} \frac{n^\alpha - 1}{n^2} < \infty$$

e quindi $S_n/n \to 0$ (sono variabili centrate).

3. L'affermazione è falsa. Se valesse la convergenza in $L^2(\mathbb{P})$, in particolare la successione sarebbe limitata, ma i calcoli del primo punto mostrano che

$$\frac{\mathbb{E}\left[S_n^2\right]}{n^2} = \frac{\sum_{k=1}^{n} k^\alpha - n}{n^2}.$$

Confrontando (stavolta dal basso) con l'integrale di x^α, si trova che

$$\liminf_{n\to\infty} \frac{\sum_{k=1}^{n} k^\alpha}{n^{\alpha+1}} > 0,$$

quindi se $\alpha > 1$ si ha

$$\frac{\mathbb{E}\left[S_n^2\right]}{n^2} \to \infty.$$

4. L'affermazione è vera. Infatti, se $\alpha > 1$,

$$\sum_n P(X_n \neq -1) = \sum_n P(X_n = n^\alpha - 1) = \sum_n \frac{1}{n^\alpha} < \infty,$$

e quindi per Borel-Cantelli (prima parte) si ha $\mathbb{P}$-q.c. che $X_n = -1$ definitivamente in n. Pertanto $\lim_{n\to\infty} X_n = 1$, $\mathbb{P}$-q.c. e quindi (teorema di Cesaro)

$$\lim_{n\to\infty} \frac{S_n}{n} = -1.$$

5. L'affermazione è falsa. Infatti, stavolta

$$\sum_n P(X_n \neq -1) = \sum_n P(X_n = n - 1) = \sum_n \frac{1}{n} = \infty,$$

quindi per Borel-Cantelli (seconda parte) si ha $\mathbb{P}$-q.c. che $X_n = n - 1$ frequentemente in n. Pertanto i termini della serie sono $X_n/n = 1 - 1/n$ frequentemente in n, in particolare non sono infinitesimi e quindi la serie non può essere convergente. Si può in alternativa osservare che le variabili X_n/n sono uniformemente limitate (da 1) e centrate, e $\operatorname{Var}(X_n/n) = \frac{n-1}{n^2}$ non è sommabile, e quindi usare uno dei teoremi visti a lezione.

CdL in Matematica, Probabilità (070AA)
A.A. 2021/22 - Prova scritta 2022-07-14

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Su uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$, siano definite $(X_k)_{k=1}^{\infty}$ variabili aleatorie reali indipendenti e identicamente distribuite, dotate di momento secondo finito e tali che $\mathbb{E}[X_i] = 0$, $\mathbb{E}\left[X_i^2\right] = 1$ per ogni k. Per ogni $n \in \mathbb{N}$, $n \geq 1$, si ponga

$$Y_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k.$$

1. Sia $(m_n)_{n=1}^{\infty}$ una successione a valori in $\mathbb{N} \setminus \{0\}$. Mostrare che $(Y_{m_n} - Y_n)_{n=1}^{\infty}$ converge verso la costante 0 in $L^2(\mathbb{P})$ se e solo se $\lim_{n\to\infty} m_n/n = 1$.
2. Mostrare che $(Y_{2n} - Y_n)_{n=1}^{\infty}$ converge in legge verso una variabile gaussiana (calcolarne i parametri).
3. Dedurre dal punto sopra che $(Y_n)_{n=1}^{\infty}$ non converge in probabilità.

Una soluzione:

1. Calcoliamo (scriviamo m invece di m_n per semplicità)

$$\mathbb{E}\left[(Y_m - Y_n)^2\right] = \mathbb{E}\left[Y_m^2\right] + \mathbb{E}\left[Y_n^2\right] - 2\mathbb{E}[Y_m Y_n].$$

È immediato vedere che

$$\mathbb{E}\left[Y_n^2\right] = \mathbb{E}\left[Y_m^2\right] = 1,$$

mentre

$$\mathbb{E}[Y_n Y_m] = \frac{1}{\sqrt{mn}} \mathbb{E}\left[\sum_{k=1}^{n}\sum_{h=1}^{m} X_k X_h\right] = \frac{1}{\sqrt{mn}} \min\{m,n\} = \min\left\{\sqrt{\frac{m}{n}}, \sqrt{\frac{n}{m}}\right\}.$$

Otteniamo quindi che

$$\mathbb{E}\left[(Y_m - Y_n)^2\right] = 2\left(1 - \min\left\{\sqrt{\frac{m}{n}}, \sqrt{\frac{n}{m}}\right\}\right).$$

Se $m_n/n \to 1$, allora anche $n/m_n \to 1$ e quindi

$$\min\left\{\sqrt{\frac{m}{n}}, \sqrt{\frac{n}{m}}\right\} \to 1,$$

da cui la convergenza verso 0 in L^2. Viceversa, se vale la convergenza verso 0, allora

$$\min\left\{\sqrt{\frac{m_n}{n}}, \sqrt{\frac{n}{m_n}}\right\} \to 1,$$

e quindi $m_n/n \to 1$ (se così non fosse, allora a meno di estrarre una sottosuccessione si avrebbe $m_n/n \to \lambda \in [0,\infty] \setminus \{1\}$ e quindi

$$\min\left\{\sqrt{\frac{m_n}{n}}, \sqrt{\frac{n}{m_n}}\right\} \to \min\{\lambda, 1/\lambda\} < 1.$$

2. Scriviamo esplicitamente la differenza come

$$Y_{2n} - Y_n = \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k + \frac{1}{\sqrt{2}\sqrt{n}}\sum_{k=n+1}^{2n} X_k.$$

Siamo interessati alla convergenza in legge, quindi basta notare che, mentre la prima somma coincide con Y_n, la seconda somma è (in legge) una copia Y_n' di Y_n da essa indipendente:

$$Y_{2n} - Y_n = \left(\frac{1}{\sqrt{2}} - 1\right) Y_n + \frac{1}{\sqrt{2}} Y_n'.$$

Per il teorema limite centrale, si ha che $Y_n \to Z$ una Gaussiana standard, e pure $Y_n' \to Z'$ una Gaussiana standard. Grazie all'indipendenza, si ha pure la convergenza in legge congiunta $(Y_n, Y_n') \to (Z, Z')$, dove Z e Z' sono indipendenti (ad esempio basta usare le funzioni caratteristiche). Infine la funzione somma è continua, quindi concludiamo che

$$Y_{2n} - Y_n \to \left(\frac{1}{\sqrt{2}} - 1\right) Z + \frac{1}{\sqrt{2}} Z',$$

che è una Gaussiana di media nulla e varianza

$$\left(\frac{1}{\sqrt{2}} - 1\right)^2 + \frac{1}{2} = 2(1 - 1/\sqrt{2}).$$

3. Per il teorema limite centrale, sappiamo che $Y_n \to Z$ in legge, dove Z è una Gaussiana standard. Se valesse la convergenza in probabilità, allora si avrebbe che $Y_{2n} \to Z$ pure (verso la stessa Z) e quindi $Y_{2n} - Y_n \to Z - Z = 0$. Ma abbiamo visto sopra che la convergenza di $Y_{2n} - Y_n$ è in legge verso una variabile Gaussiana (in particolare non costante).

Problema 2

Su uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$, siano definite due variabili aleatorie X, Y indipendenti e aventi legge uniforme (continua) sull'intervallo $[0,1]$. Si definiscano le variabili $U = \min\{X, Y\}$ e $V = \max\{X, Y\}$.

1. Mostrare che la legge della variabile (U, V) è uniforme sul triangolo
$$\left\{(u,v) \in [0,1]^2 \,:\, u \le v\right\}.$$
2. Mostrare che $\mathbb{E}[V|U] = (1-U)/2$.
3. Mostrare che, data $f : [0,1]^2 \to \mathbb{R}$ boreliana e limitata, vale l'identità
$$\mathbb{E}[f(X,Y)|U,V] = \frac{f(U,V) + f(V,U)}{2}.$$

Una soluzione:

1. Osserviamo che vale $0 \leq U \leq V \leq 1$. Dati $0 \leq u \leq v \leq 1$, calcoliamo

$$P(u \leq U, V \leq v) = P(u \leq X \leq v, u \leq Y \leq v) = (v-u)^2.$$

Da questa identità possiamo ottenere formalmente la densità, ad esempio derivando rispetto ad u, v e poi cambiando di segno:

$$f(u,v) = -\frac{\partial}{\partial u}\frac{\partial}{\partial v}(v-u)^2 = 2.$$

Abbiamo quindi ottenuto che la legge della coppia (U,V) è uniforme sul triangolo $T = \{(u,v) \in [0,1]^2 \, : u \leq v\}$. Per essere più rigorosi, basta notare che abbiamo

$$P(u \leq U, V \leq v) = \int_{[u,1]\times[0,v]\cap T} f(u,v).$$

e quindi per il criterio di coincidenza (la classe $([u,1] \times [0,v])_{0 \leq u \leq v \leq 1}$ è stabile per intersezione e genera i Boreliani di T) le due misure coincidono.

2. Avendo la densità congiunta, è una semplice applicazione della teoria vista nel corso: la densità condizionale di v data $U = u$ è uniforme nell'intervallo $[u,1]$, e quindi

$$\mathbb{E}\left[V|U=u\right] = \frac{1-u}{2}.$$

3. Il membro a destra è una funzione di U, V, quindi basta mostrare che data $\varphi(U,V)$ (limitata) si ha

$$\mathbb{E}\left[\varphi(U,V)f(X,Y)\right] = \mathbb{E}\left[\varphi(U,V)\frac{f(U,V)+f(V,U)}{2}\right].$$

Posto $A = \{U = X, V = Y\}$ e quindi $A^c = \{U = Y, V = X\}$ (a meno di un insieme trascurabile), si ha

$$\mathbb{E}\left[\varphi(U,V)f(X,Y)I_A\right] = \mathbb{E}\left[\varphi(U,V)f(U,V)I_A\right]$$

e similmente

$$\mathbb{E}\left[\varphi(U,V)f(X,Y)I_{A^c}\right] = \mathbb{E}\left[\varphi(U,V)f(V,U)I_{A^c}\right].$$

D'altronde, la legge di (U,V) non cambia se si scambiano X con Y, quindi si ha pure

$$\mathbb{E}\left[\varphi(U,V)f(U,V)I_A\right] = \mathbb{E}\left[\varphi(U,V)f(U,V)I_{A^c}\right],$$

e

$$\mathbb{E}\left[\varphi(U,V)f(V,U)I_{A^c}\right] = \mathbb{E}\left[\varphi(U,V)f(V,U)I_A\right].$$

Sommando si trova quindi la tesi.

CdL in Matematica, Probabilità (070AA)
A.A. 2021/22 - Prova scritta 2022-09-12

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Siano X e Y due variabili aleatorie indipendenti, con leggi uniformi rispettivamente sugli intervalli $[-a, a]$ e $[-b, b]$ e si ponga $Z = X + Y$.

1. Calcolare la funzione caratteristica di Z.
2. Determinare per quali $m \in \mathbb{R}$ la variabile $|Z|^m$ è integrabile *(Sugg: calcolare la densità di Z in 0.)*
3. Mostrare che
$$\int_{-\infty}^{\infty} \frac{\operatorname{sen}(at)\operatorname{sen}(bt)}{t^2} dt = \pi \min\{a, b\}.$$

Una soluzione:

1. La funzione caratteristica di una variabile uniforme su $[-a, a]$ è data per $t \neq 1$ da

$$\varphi_X(t) = \frac{1}{2a}\int_{-a}^{a} e^{itx}dx = \frac{e^{ita} - e^{-ita}}{2iat} = \frac{\sin(at)}{at}.$$

Pertanto ricordando che la funzione caratteristica della somma di variabili indipendenti è il prodotto delle due funzioni, si ottiene che (per $t \neq 1$)

$$\varphi_Z(t) = \varphi_X(t)\varphi_Y(t) = \frac{\sin(at)}{at} \cdot \frac{\sin(bt)}{bt}.$$

2. Sappiamo inoltre che $|Z|$ è limitata quasi certamente da $a+b$, quindi per ogni $m \geq 0$ $|Z|^m$ è integrabile. Per m negativo, osserviamo che il problema è la divergenza in $z = 0$. La densità di Z è data dalla formula di convoluzione

$$p(Z = z) = \int_{-\infty}^{\infty} p(X = x)p(Y = z - x)dx.$$

che in $z = 0$ dà

$$p(Z = 0) = \frac{1}{4ab}\int_{-a}^{a} I_{\{b \leq x \leq -b\}}dx = \frac{1}{4ab} 2\min\{a, b\}.$$

Inoltre, la densità è continua (si può anche utilizzare il fatto che la funzione caratteristica sia integrabile su tutto $\mathbb{R}$), quindi strettamente positiva in un intorno di 0. Pertanto, per confronto,

$$\mathbb{E}\left[|Z|^m\right] = \int_{-\infty}^{\infty} |z|^m p(Z = z)dz$$

vale $+\infty$ se $m \leq -1$, mentre è finito se $-1 < m < 0$.

3. La funzione caratteristica di Z è integrabile rispetto alla misura di Lebesgue, quindi possiamo applicare il teorema di inversione e ottenere la densità

$$p(Z = z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi_Z(t)e^{-itz}dt.$$

Posto $z = 0$, si trova l'identità

$$\frac{1}{4ab} 2 \min\{a,b\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin(at)\sin(bt)}{abt^2} dt,$$

che con

Problema 2

Dato $p \in (0,1)$, si consideri una serie infinita di lanci indipendenti di una moneta truccata che dà testa (T) con probabilità p e croce (C) con probabilità $1-p$, e si denoti con N_p minimo il numero di lanci affinché si osservino 2 teste (ad esempio, se la sequenza dei lanci è $CTCCTCTC\ldots$ si ha $N_p = 5$).

1. Calcolare la densità di N_p.
2. Dire se la famiglia delle leggi delle variabili aleatorie $(N_p)_{p\in(0,1)}$ è tesa.
3. Calcolare il limite in legge di $\frac{1}{n} N_{1/n}$ al tendere di $n \to \infty$.

Una soluzione:

1. La variabile N_p è discreta: la probabilità che $N_p = n$ coincide con la probabilità che si osservi una testa al lancio n ed esattamente una testa negli $n-1$ lanci precedenti. Poiché abbiamo quindi $n-1$ configurazioni a seconda del lancio in cui esce la prima testa, troviamo che

$$P(N_p = n) = p^2(1-p)^{n-2}(n-1).$$

2. La famiglia delle leggi non è tesa: infatti, per ogni $M \in \mathbb{N}$, si osserva che

$$\lim_{p\to 0^+} P(N_p \le M) = \sum_{n=2}^{M} \lim_{p\to 0^+} p^2(1-p)^{n-2}(n-1) = 0,$$

quindi non è possibile trovare M (dato $\varepsilon > 0$) affinché sia $P(N_p \le M) \ge 1-\varepsilon$ per ogni $p \in (0,1)$.

3. Per calcolare il limite in legge conviene calcolare la funzione caratteristica di N_p. Si trova che

$$\begin{aligned}\varphi_{N_p}(t) &= \sum_{n=2}^{\infty} p^2(1-p)^{n-2}(n-1)e^{itn} = e^{it}p^2(1-p)\sum_{n=2}^{\infty}(n-1)\left((1-p)e^{it}\right)^{n-1} \\ &= e^{it}p^2(1-p)\sum_{n=1}^{\infty} n\left((1-p)e^{it}\right)^n = \frac{e^{i2t}p^2(1-p)^2}{(1-(1-p)e^{it})^2},\end{aligned}$$

ricordando che

$$\sum_{n=1}^{\infty} nx^n = x\sum_{n=1}^{\infty} nx^{n-1} = x\frac{d}{dx}\frac{1}{1-x} = \frac{x}{(1-x)^2}.$$

Otteniamo quindi che

$$\lim_{n\to\infty} \varphi_{\frac{1}{n}N_{1/n}}(t) = \lim_{n\to\infty} \varphi_{N_{1/n}}(t/n) = \lim_{n\to\infty} \frac{1}{(n(1-(1-1/n)(1+it/n)))^2} = \frac{1}{(1-it)^2}.$$

Ricordando che $1/(1-it)$ è la funzione caratteristica di una variabile esponenziale di parametro 1, troviamo che il limite in legge è una variabile avente legge uguale alla somma di due variabili esponenziali di parametro 1 indipendenti (tale legge è la $\Gamma(2,1)$).

Problema 3

Siano X, Y, due variabili aleatorie reali indipendenti identicamente distribuite, assolutamente continue, con densità f, e si indichi con $F(t) = \int_{-\infty}^{t} f(s)ds$ la funzione di ripartizione.

1. Mostrare che la variabile aleatoria $\max\{X,Y\}$ è assolutamente continua ed esprimerne la densità in termini di f.
2. Mostrare che, per ogni $t \in \mathbb{R}$, si ha l'identità quasi certa

$$\mathbb{E}\left[I_{\{\min\{X,Y\}\leq t\}} | \max\{X,Y\}\right] = F_{\max\{X,Y\}}(t),$$

dove, per $z \in \mathbb{R}$, si pone $F_z(t) = 0$ se $F(z) = 0$, altrimenti

$$F_z(t) = \begin{cases} \frac{F(t)}{F(z)} & \text{se } t \leq z, \\ 1 & \text{se } t > z. \end{cases}$$

3. Nel caso in cui $F(t) = t^2$ per $t \in [0,1]$, calcolare esplicitamente

$$\mathbb{E}\left[\min\{X,Y\} \,|\, \max\{X,Y\}\right].$$

Una soluzione:

1. Osserviamo che la variabile $\max\{X,Y\}$ ha funzione di ripartizione

$$P(\max\{X,Y\} \leq s) = P(X \leq s)P(Y \leq s) = F(s)^2.$$

Derivando si trova che ha densità $z \mapsto 2F(z)f(z)$.

2. Date ψ e φ Borel e limitate, scriviamo

$$\begin{aligned}
\mathbb{E}\left[\psi(\min\{X,Y\})\varphi(\max\{X,Y\})\right] &= \mathbb{E}\left[I_{\{X\leq Y\}}\psi(X)\varphi(Y)\right] + \mathbb{E}\left[I_{\{Y\leq X\}}\psi(Y)\varphi(X)\right] \\
&= 2\mathbb{E}\left[I_{\{X\leq Y\}}\psi(X)\varphi(Y)\right] = 2\int_0^1 \varphi(y)f(y)\int_0^y \psi(x)f(x)dxdy \\
&= \int_0^1 \varphi(y)2F(y)f(y)\left(\int_0^y \psi(x)f(x)dx/F(y)\right)dy \\
&= \mathbb{E}\left[\varphi(\max\{X,Y\})g(\max\{X,Y\})\right],
\end{aligned}$$

dove

$$g(z) = \int_0^z \psi(x)f(x)dx/F(z),$$

(se $F(z) > 0$, altrimenti si può porre $g(z)$ ad esempio costante uguale a 0. Ne segue che

$$g(\max\{X,Y\}) = \mathbb{E}\left[\psi(\max\{X,Y\}) | \min\{X,Y\}\right].$$

Posta $\psi(x) = I_{\{x \leq t\}}$, si trova che $g(z) = F_z(t)$, così da risolvere il punto 2.

3. Usando la formula trovata sopra, poniamo ora $\psi(x) = x$ e troviamo, per $z \in [0, 1]$,

$$g(z) = \int_0^z x2xdx/z^2 = \frac{2}{3}z.$$

Segue che

$$\mathbb{E}\left[\min\{X, Y\} \,|\, \max\{X, Y\}\right] = \frac{2}{3}\max\{X, Y\}.$$