Tutorato IX
Probabilità e Statistica
a.a. 2015/2016

Argomenti: convergenza in legge; legge dei grandi numeri e teorema limite centrale; approssimazione normale; intervalli di fiducia.

Esercizio 1. **a)** Siano $\{X_n\}_n$ una successione di v.a. su $\mathbb{R}$, X una v.a. su $\mathbb{R}$ e $a, b \in \mathbb{R}$. Dimostrare che se $X_n \xrightarrow{\mathscr{L}} X$ allora $aX_n + b \xrightarrow{\mathscr{L}} aX + b$.

b) Dimostrare la seguente versione del TLC: *se $\{X_n\}_n$ è una successione di v.a. i.i.d. con media μ e varianza σ^2 allora*

$$\bar{S}_n = \frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}} \xrightarrow{\mathscr{L}} Z \sim \mathrm{N}(0, \sigma^2).$$

Esercizio 2. **a)** Sia X_n una v.a. con densità di probabilità f_n, $n = 1, 2, \ldots$ Sia f una funzione continua a tratti tale che

$$\lim_{n\to\infty} \int_{\mathbb{R}} |f_n(x) - f(x)| dx = 0.$$

Dimostrare che f è una densità di probabilità e, detta X una v.a. con densità f, $X_n \to X$ in legge.

b) Per $n = 1, 2, \ldots$, sia X_n una v.a. con f.r. $F_n(x) = x - \frac{\sin(2n\pi x)}{2n\pi}$ se $0 \le x < 1$, $F_n(x) = 0$ se $x < 0$ e $F_n(x) = 1$ se $x \ge 1$.

b1) Dimostrare che $X_n \to X$ in legge, con $X \sim \mathrm{Un}(0, 1)$. Dedurre che X ha densità, che indichiamo con f.

b2) Scrivere la densità f_n di X_n e provare che non vale la condizione del punto **a)**, cioè la successione delle densità $\{f_n\}_n$ non converge alla densità f della v.a. limite nel senso scritto in **a)**.

[Sugg.: potrebbe essere utile ricordare che se g è una funzione continua a tratti ed integrabile su $\mathbb{R}$ tale che $\int_I g(x)dx \ge 0$ per ogni intervallo I in $\mathbb{R}$ allora $g(x) \ge 0$ per quasi ogni x, cioè le x per cui non vale la disuguaglianza determinano un insieme finito o numerabile.]

Esercizio 3. Sia $\{X_n\}_n$ una successione di v.a. tali che $X_n \sim \mathrm{Un}(0, \frac{1}{n})$.

a) Studiare la convergenza in legge della successione $\{n^\gamma X_n\}_n$, $\gamma \in \mathbb{R}$.

b) Studiare la convergenza in legge di $\{(nX_n)^\beta\}_n$, $\beta \in \mathbb{R}$.

Esercizio 4. Sia $\{X_n\}_{n\ge1}$ una successione di v.a. tali che $\mathbb{P}(X_n = 0) = 1/n = 1 - \mathbb{P}(X_n = n^\alpha)$, con $\alpha \in \mathbb{R}$ fissato. Studiare la convergenza in legge di $\{X_n\}_n$.

Esercizio 5. Sia N il primo istante (aleatorio!) in cui si osserva l'uscita di testa in una serie di lanci ripetuti di una moneta. Sia $p \in (0, 1)$ la probabilità che esca testa. Studiare la convergenza in legge di $2pN$ quando $p = p_n \to 0$.

[Oss: poiché la legge di N dipende da p, se $p = p_n$ anche N dipende da n, quindi si chiede di studiare la convergenza in legge di $2p_nN_n$.]

Esercizio 6. Per $n \geq 1$, sia X_n una v.a. tale che $\mathbb{P}(X_n = n) = \alpha_n = 1 - \mathbb{P}(X_n = 0)$, dove $\{\alpha_n\}_n \subset [0,1]$.

a) Scrivere la f.c. di X_n.

b) Dimostrare che $\{X_n\}_n$ converge in legge se e solo $\lim_{n\to\infty} \alpha_n = 0$.

Esercizio 7. In una popolazione femminile (adulta) è noto che la statura media è 164 cm e che nell'88.5% dei casi la statura è inferiore a 170 cm. Si suppone che la variabile "statura" sia normalmente distribuita.

a) Calcolare la deviazione standard.

b) Consideriamo un campione di 25 donne, prese a caso nella popolazione femminile. Calcolare la probabilità che la media aritmetica delle stature osservate nel campione sia maggiore di 166 cm.

Esercizio 8. Siano X, Y v.a. indipendenti con $X \sim \mathrm{Un}(0,2)$ e $Y \sim \mathrm{Un}(-1,0)$. Sia $\{Z_k\}_k$ una successione di copie indipendenti di $Z = X - Y$ e $\{S_n\}_n$ la successione delle somme parziali: $S_n = \sum_{k=1}^n Z_k$. Dimostrare che

$$\lim_{n\to+\infty} \mathbb{P}\Big(|2S_n - 3n| < \sqrt{\frac{5n}{3}}\Big) = 2\Phi(1) - 1.$$

Esercizio 9. Sia X una v.a. con densità di probabilità

$$p(x) = \frac{3}{x^4}\,1\!\!1_{x>1}.$$

a) Dire per quali valori di β esiste $\mathbb{E}[X^\beta]$ e, se possibile, calcolare $\mathbb{E}[X]$ e $\mathrm{Var}(X)$.

Sia $\{X_n\}_n$ una successione di copie indipendenti di X e $\{S_n\}_n$ la successione delle somme parziali: $S_n = \sum_{k=1}^n X_k$.

b) Usando la disuguaglianza di Chebycev, stimare n affinché $\mathbb{P}\Big(\Big|\frac{1}{n}S_n - \frac{3}{2}\Big| > 0.1\Big) \leq 10^{-3}$.

c) Calcolare $\lim_{n\to\infty} \mathbb{P}\Big(\frac{2\,S_n - 3\,n}{\sqrt{n}} > 0\Big)$.

Esercizio 10. Un dado (non truccato) viene tirato 300 volte. Sia X la somma dei punteggi. Calcolare approssimativamente le probabilità seguenti:

a) $\mathbb{P}(X > 1000)$;

b) $\mathbb{P}(1000 < X \leq 1100)$.

[Sugg.: potrebbe essere utile ricordare che $\sum_{j=1}^k j = \frac{k(k+1)}{2}$ e $\sum_{j=1}^k j^2 = \frac{k(k+1)(2k+1)}{6}$.]

Esercizio 11. **a)** Una moneta equa viene tirata 1000 volte. Qual è la probabilità di ottenere un numero di teste compreso tra 490 e 520?

b) Una moneta equa viene tirata 1000 volte. Determinare k affinché la probabilità di ottenere un numero di teste compreso tra 490 e k sia circa 0.7.

c) Una moneta equa viene tirata più di 800 volte. Quante volte dev'essere tirata perché con probabilità maggiore di 0.6 escano più di 400 teste?

d) Trovare condizioni sulla probabilità p che in un lancio singolo esca testa affinché si abbia: $\mathbb{P}(S_{1000} \leq 201) > 0.5$

Esercizio 12. La quantità di neve, che cade al giorno, in un certo periodo dell'anno, su un tratto di autostrada, si può modellizzare come una variabile aleatoria continua, di media 1.5 pollici e deviazione standard 0.3 pollici.

a) Qual è la probabilità (approssimata) che in 50 giorni cadano meno di 70 pollici di neve?

b) Qual è la probabilità (approssimata) che in 50 giorni cadano tra i 70 e i 90 pollici di neve?

Esercizio 13. Un dado dà probabilità $p \in (0,1)$ all'uscita di 1, gli altri risultati $(2,3,4,5,6)$ risultano invece equiprobabili. Sia X il risultato di un lancio di questo dado.

a) Scrivere la legge di X.

b) Calcolare $\mu = \mathbb{E}(X)$ e $\sigma^2 = \mathrm{Var}(X)$. Trovare inoltre una stima dall'alto per $\sup_{p\in(0,1)} \sigma^2$.

Il dado viene poi lanciato n volte e, ad ogni lancio, viene registrato il risultato ottenuto.

c) Scrivere un intervallo di fiducia di livello 95% per μ e, di conseguenza, per p.

d) Vengono effettuati 1000 lanci, ottenendo una media dei risultati pari a 2.7. Cosa si può dire di p?

e) Stimare il numero di lanci n che occorre effettuare affinché l'intervallo di fiducia per p abbia ampiezza al più 10^{-2}.

[Sugg.: potrebbe essere utile ricordare che $\sum_{j=1}^{k} j = \frac{k(k+1)}{2}$ e $\sum_{j=1}^{k} j^2 = \frac{k(k+1)(2k+1)}{6}$.]

Soluzioni

Esercizio 1. a) Come al solito, se Y è una v.a. denotiamo con φ_Y la sua f.c.
Per ipotesi, sappiamo che $\varphi_{X_n}(t) \to \varphi_X(t)$ per ogni $t \in \mathbb{R}$. Poiché

$$\varphi_{aX_n+b}(t) = e^{itb}\varphi_{X_n}(at) \to e^{itb}\varphi_X(at) = \varphi_{aX+b}(t),$$

per ogni $t \in \mathbb{R}$, si ottiene $aX_n + b \xrightarrow{\mathscr{L}} aX + b$.

b) Osserviamo che $\bar{S}_n = \sigma S_n^*$. La versione classica del TLC assicura che $S_n^* \to\xrightarrow{\mathscr{L}} X \sim \mathrm{N}(0,1)$, quindi usando **a)** otteniamo che $\bar{S}_n = \sigma S_n^* \to\xrightarrow{\mathscr{L}} Z = \sigma X$. Ma se $X \sim \mathrm{N}(0,1)$ sappiamo che $Z = \sigma X \sim \mathrm{N}(0,\sigma^2)$, da cui la tesi.

Esercizio 2. a) Perché f sia una densità dev'essere $f \geq 0$ e $\int_{\mathbb{R}} f(x)\,dx = 1$.
Per ogni intervallo I di $\mathbb{R}$ si ha

$$\Big|\int_I f_n(x)\,dx - \int_I f(x)\,dx\Big| \leq \int_I |f_n(x) - f(x)|\,dx \leq \int_{\mathbb{R}} |f_n(x) - f(x)|\,dx \to 0$$

per $n \to \infty$, quindi

$$\lim_{n\to\infty} \int_I f_n(x)\,dx = \int_I f(x)\,dx, \qquad \text{per ogni intervallo } I \text{ di } \mathbb{R}.$$

Poiché f_n è una densità per ogni n, $f_n(x) \geq 0$ per quasi ogni x e quindi $\int_I f_n(x)\,dx \geq 0$ per ogni intervallo I di $\mathbb{R}$. Ma allora, $\int_I f(x)\,dx \geq 0$ per ogni intervallo I di $\mathbb{R}$, da cui segue che $f \geq 0$ per quasi ogni x. Inoltre, $\int_{\mathbb{R}} f_n(x)\,dx = 1$ per ogni n, dunque $\int_{\mathbb{R}} f(x)dx = 1$. Abbiamo quindi provato che f è una funzione integrabile su $\mathbb{R}$, quasi ovunque non negativa e tale che $\int_{\mathbb{R}} f(x)dx = 1$, dunque f è una densità di probabilità.
Siano ora X_n e X v.a. con densità, rispettivamente, f_n e f. Dette φ_n e φ le f.c. associate, si ha

$$\begin{aligned}|\varphi_n(t) - \varphi(t)| &= \Big|\int_{\mathbb{R}} e^{itx} f_n(x)dx - \int_{\mathbb{R}} e^{itx} f(x)dx\Big| = \Big|\int_{\mathbb{R}} e^{itx}\big(f_n(x) - f(x)\big)dx\Big| \\ &\leq \int_{\mathbb{R}} \Big|e^{itx}\big(f_n(x) - f(x)\big)\Big|dx = \int_{\mathbb{R}} \big|f_n(x) - f(x)\big|dx \to 0\end{aligned}$$

e quindi $X_n \to X$ in legge.

b1) Intanto, osserviamo che F_n è effettivamente una f.r.[1] Fissato $x \in \mathbb{R}$, si ha

$$\lim_{n\to\infty} F_n(x) = F(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

che è la f.r. di una v.a. $X \sim \mathrm{Un}(0,1)$, dunque $X_n \to X$ in legge e X ha densità $f(x) = 1\!\!1_{x\in(0,1)}$.

[1] Infatti, per $x \in (0,1)$, si ha

$$F_n'(x) = 1 - \cos(2n\pi x) \geq 0,$$

quindi F_n è crescente su $(0,1)$, dove si ha $0 \leq F_n(x) \leq 1$. Poiché per $x < 0$ $F_n(x) = 0 = F_n(0)$ e per $x > 1$ $F_n(x) = 1 = F_n(1)$, F_n è monotona non decrescente. Inoltre, F_n è continua per ogni n, quindi in particolare càdlàg. Infine, $F_n(-\infty) = 0$ e $F_n(+\infty) = 1$, quindi F_n è una f.r.

b2) La funzione di densità associata alla f.r. F_n è data da

$$f_n(x) = F_n'(x) = \big(1 - \cos(2n\pi x)\big) 1\!\!1_{x\in(0,1)}.$$

Si ha

$$\begin{aligned}\int_{\mathbb{R}} |f_n(x) - f(x)|dx &= \int_0^1 |\cos(2n\pi x)|dx = \frac{1}{2n\pi}\int_0^{2n\pi} |\cos t|\,dt \\ &= \frac{1}{2n\pi}\sum_{k=0}^{n-1}\int_{2k\pi}^{2(k+1)\pi} |\cos t|\,dt \\ &= \frac{1}{2n\pi}\sum_{k=0}^{n-1}\int_0^{2\pi} |\cos t|\,dt = \frac{C}{2\pi}\end{aligned}$$

dove $C = \int_0^{2\pi} |\cos t|\,dt > 0$, quindi $\int_{\mathbb{R}} |f_n(x) - f(x)|dx$ non converge a 0.

Esercizio 3 Osserviamo anzitutto che la f.r. di X_n è

$$G_n(x) = \begin{cases} 0 & \text{se } x < 0 \\ nx & \text{se } 0 \le x < 1/n \\ 1 & \text{se } x \ge 1/n \end{cases}$$

a) La f.r. di $n^\gamma X_n$ è

$$F_n^\gamma(x) = \mathbb{P}(n^\gamma X_n \le x) = G_n(x/n^\gamma) = \begin{cases} 0 & \text{se } x < 0 \\ n^{-(\gamma-1)}x & \text{se } 0 \le x < n^{\gamma-1} \\ 1 & \text{se } x \ge n^{\gamma-1} \end{cases}$$

Allora, $\lim_{n\to\infty} F_n^\gamma(x) = F^\gamma(x)$, dove $F^\gamma(x) \equiv 0$ se $\gamma > 1$ e

$$\text{se } \gamma = 1: \quad F^\gamma(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } 0 \le x < 1; \\ 1 & \text{se } x \ge 1 \end{cases}$$

$$\text{se } \gamma < 1: \quad F^\gamma(x) = \begin{cases} 0 & \text{se } x \le 0 \\ 1 & \text{se } x > 0. \end{cases}$$

Quindi se $\gamma > 1$ non c'è convergenza in legge: la funzione limite F^γ non è una f.r. perché $F^\gamma(+\infty) = 0$.
Se $\gamma = 1$, la funzione limite è la f.r. di una v.a. $X \sim \mathrm{Un}(0,1)$, dunque $nX_n \to X \sim \mathrm{Un}(0,1)$ in legge.
Se invece $\gamma < 1$, la funzione limite F^γ non è una f.r. ma lo diventa se la si modifica solo nel punto $x = 0$: posto

$$\tilde{F}^\gamma(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 & \text{se } x \ge 0 \end{cases}$$

allora $\tilde{F}^\gamma$ è la f.r. della v.a. (costante) $X = 0$ q.c. e $F_n^\gamma(x) \to \tilde{F}^\gamma(x)$ per ogni x di continuità per $\tilde{F}^\gamma$, quindi $n^\gamma X_n \to 0$ in legge.

b) Ser $\beta = 0$, $(nX_n)^\beta = 1$, quindi la convergenza è ovvia. Studiamo allora il caso $\beta \neq 1$.

La f.r. di $(n\,X_n)^\beta$ è $F_n^\beta(x) = \mathbb{P}((n\,X_n)^\beta \le x)$. Per $\beta > 0$, $F_n^\beta(x) = 0$ se $x \le 0$ e $F_n^\beta(x) = \mathbb{P}(X_n \le x^{1/\beta}/n)$ se $x > 0$, quindi

$$F_n^\beta(x) \equiv F^\beta(x) = \begin{cases} 0 & \text{se } x < 0 \\ x^{1/\beta} & \text{se } 0 \le x < 1 \\ 1 & \text{se } x \ge 1 \end{cases}$$

Per $\beta < 0$, $F_n^\beta(x) = 0$ se $x \le 0$ e se invece $x > 0$ si ha $F_n^\beta(x) = \mathbb{P}(X_n \ge x^{1/\beta}/n) = 1 - \mathbb{P}(X_n < x^{1/\beta}/n)$, quindi

$$F_n^\beta(x) \equiv F^\beta(x) = \begin{cases} 0 & \text{se } x < 1 \\ 1 - x^{1/\beta} & \text{se } x \ge 1 \end{cases}$$

Allora, per $\beta \ne 0$, F_n^β non dipende da n, quindi $(nX_n)^\beta$ converge in legge ad una v.a. X_β con f.r. F^β.

Esercizio 4 X_n assume solo due valori: 0, con probabilità $1/n$, e n^α, con probabilità $1-1/n$, È quindi semplice calcolare la f.r. F_n di X_n:

$$F_n(x) = \mathbb{P}(X_n \le x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{1}{n} & \text{se } 0 \le x < n^\alpha \\ 1 & \text{se } x \ge n^\alpha \end{cases}$$

Quindi, per ogni $x \in \mathbb{R}$ esiste $\lim_{n\to\infty} F_n(x) = G_\alpha(x)$, dove: per $\alpha < 0$,

$$G_\alpha(x) = \begin{cases} 0 & \text{se } x \le 0 \\ 1 & \text{se } x > 0; \end{cases}$$

se $\alpha = 0$,

$$G_0(x) = \begin{cases} 0 & \text{se } x < 1 \\ 1 & \text{se } x \ge 1; \end{cases}$$

se invece $\alpha > 0$,

$$G_\alpha(x) \equiv 0.$$

Allora, se $\alpha > 0$ il limite delle F_n non converge ad una f.r. (la funzione limite non diventa una f.r. neanche se la si modifica su una quantità numerabile di punti), quindi in tal caso non c'è convergenza in legge. Se invece $\alpha \le 0$ allora G_α è una f.r. ovunque eccetto che nel punto di discontinuità: dette infatti $\hat{F}_1$ e $\hat{F}_0$ la f.r. di $X = 1$ e $X = 0$ rispettivamente, si ottiene immediatamente che se $\alpha = 0$ allora $F_n(x) \to \hat{F}_1(x)$ per ogni x tale che $\Delta\hat{F}_1(x) = 0$ e che se $\alpha < 0$ allora $F_n(x) \to \hat{F}_0(x)$ per ogni x tale che $\Delta\hat{F}_0(x) = 0$. Dunque, riassumendo, possiamo dire che

- se $\alpha < 0$ si ha che $X_n \xrightarrow{\mathscr{L}} 0$;
- se $\alpha = 0$ si ha che $X_n \xrightarrow{\mathscr{L}} 1$;
- se $\alpha > 0$ la successione $\{X_n\}_n$ non converge in legge.

Esercizio 5. Come già visto tante volte, N è una v.a. discreta con legge geometrica modificata di parametro p: la densità discreta è

$$p_N(k) = (1-p)^{k-1}\,p, \quad k = 1, 2, \dots$$

Studiamo la f.c. di $2pN$:

$$\varphi_{2pN}(\theta) = \varphi_N(2p\theta),$$

dove

$$\begin{aligned}\varphi_N(t) = \mathbb{E}(e^{itN}) &= \sum_{k=1}^{\infty} e^{itk}(1-p)^{k-1}p = e^{it}p \sum_{k=1}^{\infty} \Big((1-p)e^{it}\Big)^{k-1} \\ &= e^{it}p \sum_{k=0}^{\infty} \Big((1-p)e^{it}\Big)^{k} = \frac{e^{it}p}{1-(1-p)e^{it}}.\end{aligned}$$

Se $p = p_n$, si ha $N = N_n$ e

$$\varphi_{2p_nN_n}(\theta) = \frac{p_n\, e^{2ip_n\theta}}{1-(1-p_n)e^{2ip_n\theta}}.$$

Ora, se $p_n \to 0$, $e^{2ip_n\theta} = 1 + 2ip_n\theta + o(p_n)$, quindi

$$\begin{aligned}\varphi_{2p_nN}(\theta) &= \frac{p_n(1+2ip_n\theta+o(p_n))}{1-(1-p_n)(1+2ip_n\theta+o(p_n))} \\ &= \frac{p_n(1+2ip_n\theta+o(p_n))}{-2ip_n\theta - o(p_n) + p + 2ip_n^2\theta + p_n\, o(p_n)} \\ &= \frac{1+2ip_n\theta+o(p_n)}{-2i\theta - o(p_n)/p_n + 1 + 2ip_n\theta + o(p_n)} \\ &\overset{n\to\infty}{\longrightarrow} \frac{1}{1-2i\theta} =: \varphi(\theta).\end{aligned}$$

Osserviamo che $\varphi(\theta)$ è la f.c. della legge esponenziale di parametro $1/2$. Dunque, si può concludere che $2p_nN_n$ converge in legge ad una v.a. di legge Exp$(1/2)$.

Oppure, si può anche procedere come segue. Indicando con F_n la f.r. di $2pN$, si ha $F(x) = \mathbb{P}(2pN \leq x) = \mathbb{P}(N \leq x/2p)$. Quindi, per $p = p_n$, la f.r. è $F_n(x) = 0$ se $x < 2p_n$ e per $x \geq 2p_n$,

$$\begin{aligned}F_n(x) &= \sum_{k\,:\,1\leq k\leq x/2p_n} (1-p)^{k-1}p_n = p_n \sum_{k=1}^{[x/2p_n]-1} (1-p_n)^k \\ &= p_n\,\frac{1-(1-p_n)^{[x/2p_n]}}{1-(1-p_n)} = 1-(1-p_n)^{[x/2p_n]}.\end{aligned}$$

Allora,

$$\lim_{n\to\infty} F_n(x) = F(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ 1-e^{-x/2} & \text{se } x > 0, \end{cases}$$

che è la f.r. di una Exp$(1/2)$, dunque $2p_nN_n$ converge in legge ad una v.a. di legge Exp$(1/2)$.

Esercizio 6. a) Evidentemente X_n può assumere solo i due valori n e 0, rispettivamente con probabilità α_n e $1-\alpha_n$. Dunque, la f.r. associata a X_n vale

$$F_n(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-\alpha_n & \text{se } 0 \leq x < n \\ 1 & \text{se } x > n \end{cases}$$

Invece, la f.c. vale

$$\varphi_{X_n}(t) = \mathbb{E}(e^{itX_n}) = e^{itn}\alpha_n + e^{it0}(1-\alpha_n) = e^{itn}\alpha_n + 1 - \alpha_n.$$

a) Se $\alpha_n \to 0$ allora

$$\varphi_{X_n}(t) = \mathbb{E}(e^{itX_n}) = e^{itn}\alpha_n + 1 - \alpha_n \to 1 = \varphi_X(t)$$

con $X = 0$, dunque $X_n \to 0$ in legge.
Viceversa, supponiamo che $\{X_n\}_n$ converga in legge: esiste una f.r. F tale che $F_n(x) \to F(x)$ in ogni punto x di continuità per F. Osserviamo che se $x \geq 0$ allora $F_n(x) = 1 - \alpha_n$ per ogni n tale che $n > x$, quindi $F_n(x)$ converge se e solo se $\{1-\alpha_n\}_n$ converge, il che accade se e solo se $\{\alpha_n\}_n$ è una successione convergente. Dunque, nella nostra ipotesi sappiamo che esiste α_0 tale che $\alpha_n \to \alpha_0$ per $n \to \infty$. Dimostriamo che $\alpha_0 = 0$.
Si ha,

- se $x < 0$, $F_n(x) = 0 \to 0$;
- se $x \geq 0$, $F_n(x) \to 1 - \alpha_0$.

Dunque, la f.r. limite è

$$F(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-\alpha_0 & \text{se } x \geq 0 \end{cases}$$

Ma F è una f.r. se e solo se $1 - \alpha_0 = 1$, e quindi $\alpha_0 = 0$.

Esercizio 7. a) Indicando con σ la deviazione standard di X e ricordando che $\frac{X-164}{\sigma} \sim$ N$(0,1)$, si ha

$$0.885 = \mathbb{P}(X \leq 170) = \mathbb{P}\Big(\frac{X-164}{\sigma} \leq \frac{170-164}{\sigma}\Big) = \Phi\Big(\frac{6}{\sigma}\Big).$$

Ora, $0.885 \simeq \Phi^{-1}(1.2)$, quindi

$$\frac{6}{\sigma} = 1.2, \quad \text{cioè} \quad \sigma = 5.$$

Ciò significa che mediamente la statura delle donne è 164 cm, con un errore in gran parte compreso tra $+5$ e -5 cm, ovvero per lo più la statura è compresa tra 159 cm e 169 cm:

$$\mathbb{P}(159 \leq X \leq 169) = \mathbb{P}\Big(-1 \leq \frac{X-164}{5} \leq 1\Big) = 2\Phi(1) - 1 = 0.6826.$$

b) Siano $X_1, \ldots, X_{25}$ le stature osservate nel campione, che possiamo supporre indipendenti. La media aritmetica è $\overline{X}_{25} = \frac{1}{25}\sum_{i=1}^{25} X_i$. Ora, poiché $X_i \sim$ N$(164, 25)$ per ogni i, sappiamo che $\overline{X}_{25} \sim$ N$(164, 1)$, quindi $\overline{X}_{25} - 164 \sim$ N$(0,1)$. Allora,

$$\mathbb{P}(\overline{X}_{25} \leq 166) = \mathbb{P}\Big(\overline{X}_{25} - 164 \leq 166 - 164\Big) = \Phi(2) = 0.99725.$$

Esercizio 8. a) Ricordando che la media e la varianza di una Un$(0,1)$ sono 1/2 e 1/12 rispettivamente si ha subito

$$\mathbb{E}[X] = 1 \quad \text{Var}(X) = \frac{4}{12} = \frac{1}{3}$$

$$\mathbb{E}[Y] = -\frac{1}{2} \quad \mathrm{Var}(Y) = \frac{1}{12}.$$

Pertanto,

$$\mathbb{E}[Z] = \mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y] = \frac{3}{2}$$

e dato che X ed Y sono indipendenti

$$\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{5}{12},$$

da cui si ottiene,

$$\mathbb{E}[S_n] = \frac{3}{2}n, \quad \mathrm{Var}(S_n) = \frac{5}{12}n.$$

Quindi per il TLC la legge della v.a. $\frac{S_n - \frac{3}{2}n}{\sqrt{\frac{5}{12}n}} = \frac{2S_n - 3n}{\sqrt{\frac{5}{3}n}}$ per $n \to +\infty$ tende alla legge di una $\mathrm{N}(0,1)$. Indicando con $Z^* \sim \mathrm{N}(0,1)$,

$$\begin{aligned}
\lim_{n\to+\infty} \mathbb{P}\Big(|2S_n - 3n| > \sqrt{\frac{5n}{3}}\Big) &= \lim_{n\to+\infty} \mathbb{P}\Big(\frac{|2S_n - 3n|}{\sqrt{\frac{5}{3}n}} > 1\Big) \\
&= \lim_{n\to+\infty} \mathbb{P}(|Z^*| > 1) = 2\Phi(1) - 1.
\end{aligned}$$

Esercizio 9. a) Perché esista $\mathbb{E}(X^\beta)$ deve esistere $\int_{\mathbb{R}} |x|^\beta p(x)\,dx$. Ora, $p(x) \equiv 0$ per $x \leq 1$, basta quindi lavorare con $\int_1^{+\infty} x^\beta p(x)\,dx$. Il problema, se c'è un problema, qui è ovviamente all'infinito: studiamo il limite, per $\alpha \to +\infty$, di $\int_1^\alpha x^\beta p(x)\,dx$:

$$\int_1^\alpha x^\beta p(x)\,dx = 3\int_1^\alpha x^{\beta-4}\,dx = \begin{cases} 3\,\frac{x^{\beta-3}}{\beta-3}\,\Big|_1^\alpha = \frac{3}{3-\beta}(1 - \alpha^{\beta-3}) & \text{se } \beta - 4 \neq -1 \\ 3\ln x\,\Big|_1^\alpha = 3\ln\alpha & \text{se } \beta - 4 = -1 \end{cases}$$

Quindi, esiste finito il limite per $\alpha \to +\infty$ di $\int_1^\alpha x^\beta p(x)\,dx$ se e solo se $\beta - 3 < 0$, cioè $\beta < 3$. In particolare quindi l'integrale $\int_{\mathbb{R}} |x|^\beta p(x)\,dx$ esiste per $\beta = 1, 2$, dunque esistono sia media che varianza di X e

$$\begin{aligned}
\mathbb{E}(X) &= \int_1^{+\infty} x\,p(x)\,dx = \frac{3}{2}, \\
\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) &= \int_1^{+\infty} x^2\,p(x)\,dx - \frac{9}{4} = 3 - \frac{9}{4} = \frac{3}{4}.
\end{aligned}$$

b) Poiché $\frac{3}{2} = \mathbb{E}(X_i)$ per ogni i, allora $\frac{3}{2} = \mathbb{E}(\frac{1}{n}S_n)$, quindi la disuguaglianza di Chebycev si può effettivamente usare e

$$\mathbb{P}\Big(\Big|\frac{1}{n}S_n - \frac{3}{2}\Big| > 0.1\Big) \leq \frac{\mathrm{Var}\Big(\frac{1}{n}\,S_n\Big)}{0.1^2} = 100\,\frac{1}{n^2}\cdot n\,\mathrm{Var}(X_i) = \frac{300}{4\,n}.$$

Quindi, perché $\mathbb{P}\Big(\Big|\frac{1}{n}S_n - \frac{3}{2}\Big| > 0.1\Big) \leq 10^{-3}$ basta che $\frac{300}{4\,n} \leq 10^{-3}$, cioè $n \geq 13334$.

c) Osserviamo che

$$\frac{2\,S_n - 3\,n}{\sqrt{n}} = 2\,\frac{S_n - \frac{3}{2}n}{\sqrt{n}} = 2\sqrt{\frac{3}{4}}\,\frac{S_n - \frac{3}{2}n}{\sqrt{3n/4}}$$

e, dal TLC, la legge di $\frac{S_n - \frac{3}{2}n}{\sqrt{3n/4}}$ converge, per $n \to +\infty$ ad una gaussiana standard. Allora,

$$\lim_{n\to\infty} \mathbb{P}\Big(\frac{2\,S_n - 3\,n}{\sqrt{n}} > 0\Big) = \lim_{n\to\infty} \mathbb{P}\Big(\sqrt{3}\,\frac{S_n - \frac{3}{2}n}{\sqrt{3n/4}} > 0\Big) = \lim_{n\to\infty} \mathbb{P}\Big(\frac{S_n - \frac{3}{2}n}{\sqrt{3n/4}} > 0\Big)$$
$$= 1 - \Phi(0) = 1 - \frac{1}{2} = 0.5.$$

Esercizio 10. Siano $X_1, X_2, \dots, X_{300}$ i risultati ottenuti nei 300 lanci. Le v.a. X_i sono indipendenti ed identicamente distribuite, uniformi su $\{1,2,3,4,5,6\}$. Calcoliamo la media e lo scarto quadratico medio delle X_i.

$$\mu = \mathbb{E}(X_i) = \frac{1}{6}\sum_{j=1} j = \frac{1}{6}\cdot\frac{6\cdot 7}{2} = \frac{7}{2}$$

mentre

$$\mathbb{E}(X_i^2) = \frac{1}{6}\sum_{j=1} j^2 = \frac{1}{6}\cdot\frac{6\cdot 7\cdot 13}{6} = \frac{91}{6}$$

da cui

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{\mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2} = \sqrt{\frac{91}{6} - \Big(\frac{7}{2}\Big)^2} = \sqrt{\frac{35}{12}} = 1.71.$$

Se $X = X_1 + X_2 + \ldots + X_{300}$ è la somma dei punteggi, dal TLC si ha,

$$\frac{X - n\mu}{\sigma\sqrt{n}} = \frac{X - 1050}{29.58} \sim \mathrm{N}(0,1).$$

a) Si ha

$$\begin{aligned}\mathbb{P}(X > 1000) &= \mathbb{P}\Big(\frac{X - 1050}{29.58} > \frac{1000 - 1050}{29.58}\Big)\\ &\approx \mathbb{P}(Z^* > -1.69) = \Phi(1.69) = 0.9545,\end{aligned}$$

dove $Z^* \sim \mathrm{N}(0,1)$ e, come sempre, Φ è la f.r. di una normale standard.
Usando la correzione di continuità, indicando con Z una gaussiana con stessa media e stessa varianza di X, avremmo avuto,

$$\begin{aligned}\mathbb{P}(X > 1000) &\approx \mathbb{P}(Z > 1000.5) = \mathbb{P}\Big(\frac{Z - 1050}{29.58} > \frac{1000.5 - 1050}{29.58}\Big)\\ &= \mathbb{P}(Z^* > -1.67) = \Phi(1.67) = 0.9525.\end{aligned}$$

b) Si ha

$$\begin{aligned}\mathbb{P}(1000 < X \le 1100) &= \mathbb{P}\Big(\frac{1000 - 1050}{29.58} < \frac{X - 1050}{29.58} \le \frac{1100 - 1050}{29.58}\Big)\\ &\approx \mathbb{P}(-1.69 < Z^* \le 1.69) = \Phi(1.69) - \Phi(-1.69)\\ &= 2\cdot\Phi(1.69) - 1 = 0.9090.\end{aligned}$$

Usando la correzione di continuità, indicando con Z una gaussiana con stessa media e stessa varianza di X, avremmo avuto,

$$\begin{aligned}\mathbb{P}(1000 < X \le 1100) &\approx \mathbb{P}(1000.5 < Z \le 1100.5)\\ &= \mathbb{P}\Big(\frac{1000.5 - 1050}{29.58} < \frac{Z - 1050}{29.58} \le \frac{1100.5 - 1050}{29.58}\Big)\\ &= \mathbb{P}(-1.67 < Z^* \le 1.71) = \Phi(1.71) - \Phi(-1.67) = 0.9079.\end{aligned}$$

Esercizio 11. a) Qui $p = \frac{1}{2}$ e $n = 1000$. Si chiede $\mathbb{P}(490 \le S_{1000} \le 510)$. Indicando con Z una normale con stessa media e stessa varianza di S_{1000} e con Z^* una normale standard, si ha

$$\begin{aligned}
\mathbb{P}(490 \le S_{1000} \le 520) &= \mathbb{P}\Big(\frac{489.5-500}{\sqrt{250}} < \frac{S_{1000}-500}{\sqrt{250}} \le \frac{520.5-500}{\sqrt{250}}\Big) \\
&\simeq \mathbb{P}\Big(\frac{489.5-500}{\sqrt{250}} < Z^* \le \frac{520.5-500}{\sqrt{250}}\Big) \\
&= \mathbb{P}(-0.6641 \le Z^* \le 1.2965) = \Phi(1.2965) - \Phi(-0.6641) \\
&= \Phi(1.2965) - 1 + \Phi(0.6641) \\
&= 0.90147 - 1 + 0.7454 = 0.64687,
\end{aligned}$$

quindi $\mathbb{P}(490 \le S_{1000} \le 520) \simeq 0.64687$.

b) Qui $p = \frac{1}{2}$ e $n = 1000$. Si chiede k affinché $\mathbb{P}(490 \le S_{1000} \le k) = 0.7$. Indicando con Z una normale con stessa media e stessa varianza di S_{1000} e con Z^* una normale standard, si ha

$$\begin{aligned}
\mathbb{P}(490 \le S_{1000} \le k) &\simeq \mathbb{P}(489.5 < Z \le k + 0.5) \\
&= \mathbb{P}\Big(\frac{489.5-500}{\sqrt{250}} < Z^* \le \frac{k+0.5-500}{\sqrt{250}}\Big) \\
&= \mathbb{P}\Big(-0.6641 \le Z^* \le \frac{k+0.5-500}{\sqrt{250}}\Big) \\
&= \Phi\Big(\frac{k+0.5-500}{\sqrt{250}}\Big) - \Phi(-0.6641) = 0.7,
\end{aligned}$$

Allora

$$\Phi\Big(\frac{k+0.5-500}{\sqrt{250}}\Big) = 0.7 + \Phi(-0.6641) = 0.7 + 1 - \Phi(0.6641) = 0.9546.$$

Ora, uno sguardo alle tavole dà $\phi_{0.9546} = \Phi^{-1}(0.9546) \simeq 1.7$, quindi

$$\frac{k+0.5-500}{\sqrt{250}} \simeq 1.7, \qquad \text{cioè} \qquad k \simeq 499.5 + 1.7 \cdot \sqrt{250} = 526.3794.$$

c) Qui $p = \frac{1}{2}$. Si chiede $n > 800$ affinché $\mathbb{P}(S_n > 400) > 0.6$. Indicando con Z una normale con stessa media e stessa varianza di S_n e con Z^* una normale standard, si ha

$$\mathbb{P}(S_n > 400) \simeq \mathbb{P}(Z > 400.5) = \mathbb{P}\Big(Z^* > \frac{400.5 - n/2}{\sqrt{n}/2}\Big) = 1 - \Phi\Big(\frac{400.5 - n/2}{\sqrt{n}/2}\Big) > 0.6.$$

Basta allora cercare n affinché

$$\Phi\Big(\frac{801-n}{\sqrt{n}}\Big) < 0.4$$

Ora, poiché $n > 800$, $\frac{801-n}{\sqrt{n}} \le 0$, quindi

$$\Phi\Big(\frac{801-n}{\sqrt{n}}\Big) = 1 - \Phi\Big(\frac{n-801}{\sqrt{n}}\Big) < 0.4$$

cioè
$$\Phi\Big(\frac{n-801}{\sqrt{n}}\Big) > 0.6.$$
Ma $\Phi^{-1}(0.6) \simeq 0.26$ e poiché Φ è crescente possiamo scrivere
$$\frac{n-801}{\sqrt{n}} > 0.26.$$
Risolvendo la disequazione, otteniamo $n > 808.5$, quindi basta tirare la moneta 809 volte.

d) Qui $n = 1000$ e si cerca p affinché $\mathbb{P}(S_n < 200) > 0.5$. Indicando con Z^* una normale standard, si ha
$$\begin{aligned}\mathbb{P}(S_{1000} \le 201) &\simeq \mathbb{P}\Big(Z^* < \frac{201.5 - 1000\,p}{\sqrt{1000\,p(1-p)}}\Big)\\ &= \Phi\Big(\frac{201.5 - 1000\,p}{\sqrt{1000\,p(1-p)}}\Big) > 0.5.\end{aligned}$$
Ma allora $201.5 - 1000\,p > 0$, cioè $p < 0.2015$.

Esercizio 12. Siano X_i, $i = 1, 2, \ldots, 50$, le quantità di neve che cadono in ciascun giorno. Allora $\mu = \mathbb{E}[X_i] = 1.5$ e $\sigma = \sqrt{\mathrm{Var}(X_i)} = 0.3$. La quantità di neve che cade in 50 giorni è allora $X = X_1 + X_2 + \ldots + X_{50}$. Supponendo che la quantità di neve che cade in ciascun giorno è indipendente da quella che cade negli altri giorni, possiamo applicare il TLC e quindi
$$\frac{X - 50\mu}{\sigma\sqrt{50}} = \frac{X - 75}{2.12} \sim \mathrm{N}(0, 1).$$
a) La probabilità richiesta è $\mathbb{P}(X < 70)$. Si ha
$$\begin{aligned}\mathbb{P}(X < 70) &= \mathbb{P}\Big(\frac{X-75}{2.12} < \frac{70-75}{2.12}\Big)\\ &\approx \mathbb{P}(Z^* < -2.36) = \Phi(-2.36) = 0.0091,\end{aligned}$$
b) La probabilità richiesta è $\mathbb{P}(70 < X < 90)$.
$$\begin{aligned}\mathbb{P}(70 < X < 90) &= \mathbb{P}\Big(\frac{70-75}{2.12} < \frac{X-75}{2.12} < \frac{90-75}{2.12}\Big)\\ &\approx \mathbb{P}(-2.36 < Z^* \le 7.08) = \Phi(7.08) - \Phi(-2.36) = 0.9909.\end{aligned}$$
ATTENZIONE! In questo esercizio non si applica la correzione di continuità. Le variabili X_i, quantità di neve, sono esse stesse continue, la correzione di continuità si applica nel passaggio dal discreto al continuo, quindi per v.a. discrete.

Esercizio 13. a) Calcoliamo la probabilità di ottenere un risultato diverso da 1. Sappiamo che un risultato diverso da 1 avviene con la stessa probabilità, chiamamola α. Detta p_X la densità discreta di X, si ha $p_X(1) = p$, $p_X(j) = \alpha$ per $j \in \{2, 3, 4, 5, 6\}$ e $p_X = 0$ altrimenti. Poiché $\sum_j p_X(j) = 1$, si ha
$$1 = p + \sum_{j=2}^{6} \alpha = p + 5\alpha$$

da cui segue che $\alpha = \frac{1-p}{5}$. Quindi,

$$p_X(1) = p, \text{ per } j \in \{2, \dots, 6\} \text{ si ha } p_X(j) = \tfrac{1-p}{5} \text{ e } p_X = 0 \text{ altrimenti.}$$

b) Si ha

$$\mu = \mathbb{E}(X) = p + \sum_{j=2}^{6} j \frac{1-p}{5} = p + \frac{1-p}{5}\Big(\frac{6 \cdot 7}{2} - 1\Big) = p + 4(1-p) = 4 - 3p.$$

Per la varianza, calcoliamo dapprima $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = p + \sum_{j=2}^{6} j^2 \frac{1-p}{5} = p + \frac{1-p}{5}\Big(\frac{6 \cdot 7 \cdot 13}{6} - 1\Big) = p + 18(1-p) = 18 - 17p.$$

Quindi

$$\sigma^2 = \mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 18 - 17p - (4-3p)^2 = 2 + 7p - 9p^2.$$

Osserviamo che $[0,1] \ni p \mapsto \sigma^2 = \sigma^2(p)$ è una parabola con concavità verso il basso, quindi il max si ottiene annullando la derivata prima:

$$\frac{d}{dp}\sigma^2 = 7 - 18p = 0 \text{ se e solo se } p = \frac{7}{18}.$$

Dunque,

$$\sup_{p \in (0,1)} \sigma^2 = \max_{p \in [0,1]} \sigma^2 = 2 + 7p - 9p^2\Big|_{p=\frac{7}{18}} = \frac{121}{36}.$$

c) Indichiamo con X_k il risultato ottenuto al k-esimo lancio, $k = 1, 2, \dots, n$, e con $\bar{X}_n$ la media empirica di risultati ottenuti. Un IF al 5% per μ ha estremi

$$\bar{X}_n \pm \frac{\sigma}{\sqrt{n}} 1.96.$$

Poiché σ non è noto, possiamo usare la stima in **b)**: $\sigma \leq \sqrt{\frac{121}{36}} = \frac{11}{6}$, quindi possiamo prendere come IF per μ l'intervallo di estremi

$$\bar{X}_n \pm \frac{11}{6\sqrt{n}} 1.96.$$

Per avere un IF per p, basta usare quello per μ e riscriverlo in termini di p: $\mu = 4 - 3p$, dunque

$$\bar{X}_n - \frac{11}{6\sqrt{n}} 1.96 \leq \mu \leq \bar{X}_n + \frac{11}{6\sqrt{n}} 1.96$$
$$\Updownarrow$$
$$\bar{X}_n - \frac{11}{6\sqrt{n}} 1.96 \leq 4 - 3p \leq \bar{X}_n + \frac{11}{6\sqrt{n}} 1.96$$
$$\Updownarrow$$
$$\frac{4 - \bar{X}_n}{3} - \frac{11}{18\sqrt{n}} 1.96 \leq p \leq \frac{4 - \bar{X}_n}{3} + \frac{11}{18\sqrt{n}} 1.96$$

Infatti, si ha

$$\mathbb{P}\Big(\Big|p - \frac{4-\bar{X}_n}{3}\Big| \le \frac{11}{18\sqrt{n}}\,1.96\Big) = \mathbb{P}\Big(\big|\mu - \bar{X}_n\big| \le \frac{11}{6\sqrt{n}}\,1.96\Big)$$
$$\ge \mathbb{P}\Big(\big|\mu - \bar{X}_n\big| \le \frac{\sigma}{\sqrt{n}}\,1.96\Big) \simeq 0.95$$

d) Se $n = 1000$ e si osserva $\bar{X}_n = 2.7$, allora $\frac{11}{18\sqrt{n}}\,1.96 = 0.039$ e $\frac{4-\bar{X}_n}{3} = 0.43$. Quindi, l'IF osservato per p è

$$[0.43 - 0.039, 0.43 + 0.039] = [0.391, 0.469],$$

dunque possiamo ragionevolmente pensare che l'esatto valore di p cada in questo intervallo. Osserviamo però che l'ampiezza è pari a 0.078, che sembra un po' troppo grande se si vuole stimare con precisione un numero (qui, p) in $(0, 1)$. Insomma, se si vuole essere precisi, occorrerebbe fare qualche lancio in più.

e) L'ampiezza dell'IF che abbiamo trovato è pari $2 \times \frac{11}{6\sqrt{n}}\,1.96 = \frac{11}{3\sqrt{n}}\,1.96$. Basta quindi trovare n affinché

$$\frac{11}{3\sqrt{n}}\,1.96 \le 10^{-2},$$

cioè

$$\sqrt{n} \ge \frac{11}{3 \cdot 10^{-2}}\,1.96 = 718.6 \quad \text{e quindi} \quad n \ge 718.6^2 = 516481.7.$$

Basterà quindi prendere $n \ge 516482$ (che è ben più grande di 1000!!).