## 0.1 Teorema limite centrale

**Exercise 1** *Siano $X_1, ..., X_n$ v.a. indipendenti, identicamente distribuite, con varianza finita $\sigma^2$ e media $\mu$. Allora*

$$Z_n := \frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}}$$

*ha media zero e varianza uno.*

Il teorema limite centrale vale sotto diverse ipotesi. La seguente versione porta il nome di Teorema di P. Lévy (o Lindeberg-Lévy). Nel caso particolare in cui le v.a. $X_n$ siano delle Bernoulli, esso porta il nome di Teorema di De Moivre-Laplace. e si può dimostrare per via combinatorica.

**Theorem 2** *Sia $(X_n)$ una successione di v.a. indipendenti, identicamente distribuite, con varianza finita $\sigma^2$ e media $\mu$. Allora la v.a.*

$$Z_n := \frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}}$$

*converge in legge ad una gaussiana canonica $N(0,1)$. In altre parole, per ogni $a < b$ vale*

$$\lim_{n\to\infty} P\left(a \leq \frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}} \leq b\right) = \Phi(b) - \Phi(a)$$

*dove $\Phi$ indica la cdf normale standard.*

Prima di procedere alla dimostrazione, osserviamo che, in base all'esercizio preposto al teorema, la v.a. $Z_n$ ha media zero e varianza uno. Però non è in generale gaussiana e non è ovvio che lo diventi al limite per $n \to \infty$. Questa è la parte difficile del teorema.

**Proof.** Calcoliamo la funzione generatrice $\varphi_n(t)$ di $Z_n$ e mostriamo che, per ogni $t$, essa converge a $e^{-t^2/2}$. Questo implica la convergenza in legge di $Z_n$ alla $N(0,1)$.

Osserviamo che

$$Z_n = \frac{\frac{X_1-\mu}{\sigma} + ... + \frac{X_n-\mu}{\sigma}}{\sqrt{n}}$$

dove le v.a. $Y_n = \frac{X_n-\mu}{\sigma}$ sono indipendenti ed hanno media zero e varianza uno. Quindi basta dimostrare il teorema in questo caso.

Supponiamo allora $\mu = 0$, $\sigma = 1$. Abbiamo

$$\varphi_n(t) = \varphi_{X_1+...+X_n}\left(\frac{t}{\sqrt{n}}\right) = \varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n.$$

Allora, usando lo sviluppo di Taylor di $\varphi_{X_1}(t)$ ed il fatto che $E[X_1] = 0$ e $E[X_1^2] = 1$, vale

$$\varphi_{X_1}(t) = 1 + \frac{t^2}{2} + o\left(t^2\right).$$

Quindi

$$\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right).$$

Pertanto vale

$$\varphi_n(t) = \left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n.$$

Passando ai logaritmi abbiamo

$$\log \varphi_n(t) = n \log\left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)$$

ed usando il limite notevole $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$ si ottiene ($t$ è fissato)

$$\begin{aligned}\lim_{n\to\infty} \log \varphi_n(t) &= \lim_{n\to\infty} n\left(\frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right) \frac{\log\left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)}{\frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)} \\ &= \lim_{n\to\infty} n\left(\frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right) = \frac{t^2}{2}.\end{aligned}$$

Quindi

$$\lim_{n\to\infty} \varphi_n(t) = e^{\frac{t^2}{2}}.$$

La dimostrazione è completa. ■

Il caso particolare in cui le v.a. $X_n$ sono Bernoulli $B(1,p)$ è particolarmente rilevante. In questo caso $S_n := X_1 + ... + X_n$ è una binomiale $B(n,p)$ ed il teorema dice che

$$Z_n := \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

converge in legge ad una gaussiana canonica $N(0,1)$. E' un teorema di *convergenza della binomiale alla gaussiana*, che si affianca al teorema degli eventi rari. Qui, per vedere la convergenza, bisogna standardizzare la binomiale $S_n$ ($Z_n$ è la sua standardizzata).

Può sembrare assurdo che le binomiali approssimino contemporaneamente sia le Poisson sia le gaussiane. In effetti il regime limite è molto diverso: nel teorema degli eventi rari $p$ non è fissato, tende a zero come $\frac{\lambda}{n}$, mentre nel teorema limite centrale è fissato. Se però non si considera il limite vero e proprio ma solo l'approssimazione per valori grandi o piccoli dei parametri in gioco, ci sono effettivamente delle situazioni in cui le due approssimazioni sono abbastanza buone entrambe e si sovrappongono un po'. A parte questo, il consiglio è di usare il teorema degli eventi rari quando il prodotto $np$ è un numero dell'ordine dell'unità (es. 5), ed $n$ è ovviamente non troppo piccolo (es. 20, 30). Se ad esempio $np = 3$ e $n = 30$, allora $p = \frac{3}{30} = 0.1$, è piuttosto piccolo. In queste situazioni l'uso del teorema limite centrale non prodice risultati molto precisi. Meglio che $p$ sia più "interno" all'intervallo $(0,1)$, non così estremo, per una buona applicazione del TLC (ma se $n$ è più grande, allora si possono accettare $p$ più piccoli).

Infine, sempre nell'ambito dell'approssimazione gaussiana della binomiale, se si vuole un risultato più preciso conviene usare la *correzione di continuità*. Supponiamo di dover calcolare $P(S_n \leq 25)$. Siccome $S_n$ assume valori solo negli interi, questo è uguale a $P(S_n < 26)$. Le due approssimazioni darebbero

$$P(S_n \leq 25) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{25 - n\mu}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{25 - n\mu}{\sigma\sqrt{n}}\right)$$
$$P(S_n < 26) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{26 - n\mu}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{26 - n\mu}{\sigma\sqrt{n}}\right)$$

per cui in genere si ottiene un risultato più preciso prendendo

$$P(S_n \leq 25) \approx \Phi\left(\frac{25.5 - n\mu}{\sigma\sqrt{n}}\right).$$