Marco Minozzo e Annamaria Guolo
Laurea in Economia del Commercio Internazionale
Laurea in Economia e Amministrazione delle Imprese
Università degli Studi di Verona (sede di Vicenza)

Calcolo delle Probabilità
Soluzioni 13. Legge (debole) dei grandi numeri

Combinazioni lineari di v.a.

Valore atteso di una combinazione lineare. Siano $X_1, \ldots, X_n$, n variabili aleatorie definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathsf{P})$ e siano $c_1, \ldots, c_n$ costanti reali. Allora

$$\mathrm{E}\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i \mathrm{E}(X_i).$$

Varianza di una combinazione lineare. Siano $X_1, \ldots, X_n$, n variabili aleatorie definite sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathsf{P})$ e siano $c_1, \ldots, c_n$ costanti reali. Allora

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^n c_i X_i\right) &= \sum_{i=1}^n c_i^2 \mathrm{Var}(X_i) + \sum_{i=1}^n \sum_{\substack{j=1\\ j\neq i}}^n c_i c_j \mathrm{Cov}(X_i, X_j)\\ &= \sum_{i=1}^n c_i^2 \mathrm{Var}(X_i) + 2\sum_{i=1}^n \sum_{\substack{j=1\\ j< i}}^n c_i c_j \mathrm{Cov}(X_i, X_j).\end{aligned}$$

Nel caso in cui le variabili aleatorie $X_1, \ldots, X_n$ siano a due a due incorrelate, allora

$$\mathrm{Var}\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i^2 \mathrm{Var}(X_i).$$

Correlazione nel caso di più di due variabili. Date due variabili aleatorie X ed Y si ha che $-1 \leq \mathrm{Corr}(X, Y) \leq 1$. Date più di due variabili aleatorie, ad esempio, X, Y e Z, i valori che possono assumere le correlazioni $\mathrm{Corr}(X, Y)$, $\mathrm{Corr}(X, Z)$ e $\mathrm{Corr}(Y, Z)$ sono tra loro vincolati, e non possono assumere liberamente l'uno dagli altri qualsiasi valore tra -1 e 1.

Esempio. Siano date tre variabili aleatorie dicotomiche X, Y e Z i cui possibili valori sono solamente -1 e 1. Allora, se $\mathrm{Corr}(X, Y) = -1$ e $\mathrm{Corr}(X, Z) = -1$, abbiamo che quando $X = 1$, si ha $Y = Z = -1$, e quando $X = -1$, si ha $Y = Z = 1$. Segue che Y e Z assumono sempre lo stesso valore e perciò $\mathrm{Corr}(Y, Z)$ può solo essere pari a 1.

Media campionaria. Siano $X_1, \ldots, X_n$ variabili aleatorie indipendenti con valore atteso pari a μ e varianza pari a σ^2 (non necessariamente identicamente distribuite) e sia $\overline{X}_n = (1/n)\sum_{i=1}^n X_i$. Allora

$$\mathrm{E}(\overline{X}_n) = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n}\sum_{i=1}^n \mathrm{E}(X_i) = \frac{1}{n} n\mu = \mu,$$

e

$$\mathrm{Var}(\overline{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

Legge (debole) dei grandi numeri. Nella formulazione più semplice, la legge dei grandi numeri riassume il risultato sperimentale secondo il quale, al crescere delle ripetizioni di un esperimento sotto identiche condizioni, la frequenza relativa di un evento tende a stabilizzarsi.
In particolare, siano $X_1, \ldots, X_n$ v.a. indipendenti ed identicamente distribuite con valore atteso μ. Allora, la media aritmetica $\sum_{i=1}^{n} X_i/n$ tende, al crescere di n, alla quantità costante μ.

Combinazioni lineari

Esercizio A. a) Assumendo $\text{Cov}(X_1, X_2) = 0,5$, valore atteso e varianza di Y sono dati da

$$\text{E}(Y) = 2 \cdot 1 + 1 \cdot 2 = 4 \quad \text{e} \quad \text{Var}(Y) = 4 \cdot 1 + 1 \cdot 3 + 2 \cdot 2 \cdot 1 \cdot 0,5 = 9.$$

b) Assumendo X_1 ed X_2 indipendenti, valore atteso e varianza di Y sono dati da

$$\text{E}(Y) = 2 \cdot 1 + 1 \cdot 2 = 4 \quad \text{e} \quad \text{Var}(Y) = 4 \cdot 1 + 1 \cdot 3 = 7.$$

Esercizio B. a) Indicando con $X = 10X_1 + 20X_2 + 30X_3$ il valore giornaliero del portafoglio, si ha:

$$\begin{aligned}
\text{E}(X) &= 10 \cdot 3 + 20 \cdot 2,5 + 30 \cdot 4 = 200, \\
\text{Var}(X) &= 10^2 \cdot 0,5 + 20^2 \cdot 0,8 + 30^2 \cdot 1 + 2 \cdot 10 \cdot 20 \cdot 0,6 + 2 \cdot 10 \cdot 30 \cdot 0,6 + 2 \cdot 20 \cdot 30 \cdot 0,85 \\
&= 2890.
\end{aligned}$$

b) Nel caso in cui i tre titoli si comportino in modo indipendente si ottiene:

$$\begin{aligned}
\text{E}(X) &= 10 \cdot 3 + 20 \cdot 2,5 + 30 \cdot 4 = 200, \\
\text{Var}(X) &= 10^2 \cdot 0,5 + 20^2 \cdot 0,8 + 30^2 \cdot 1 = 1270.
\end{aligned}$$

Legge (debole) dei grandi numeri

Esercizio C. Per prima cosa verifichiamo che $\text{E}(X) = \mu$ e $\text{Var}(X) = 1,2$. Infatti:

$$\begin{aligned}
\text{E}(X) &= (1/10) \cdot (\mu - 2) + (2/10) \cdot (\mu - 1) + (4/10) \cdot (\mu) + (2/10) \cdot (\mu + 1) + (1/10) \cdot (\mu + 2) \\
&= [(1/10) + (2/10) + (4/10) + (2/10) + (1/10)] \cdot \mu - (2/10) - (2/10) + (2/10) + (2/10) \\
&= \mu, \\
\text{Var}(X) &= \text{E}[(X - \mu)^2] = (1/10) \cdot (-2)^2 + (2/10) \cdot (-1)^2 + (4/10) \cdot 0 + (2/10) \cdot (1)^2 + (1/10) \cdot (2)^2 \\
&= (4/10) + (2/10) + (2/10) + (4/10) \\
&= 12/10 = 1,2.
\end{aligned}$$

a) Valore atteso e varianza di $\overline{X}_n$ sono dati da:

$$\text{E}(\overline{X}_n) = \text{E}(X) = \mu, \quad \text{e} \quad \text{Var}(\overline{X}_n) = \frac{\text{Var}(X)}{n} = \frac{1,2}{n}.$$

b) Applicando la disuguaglianza di Tcebycheff si ottiene:

$$\mathsf{P}(|\overline{X}_{100} - \mu| < 0,5) = \mathsf{P}\Big(|\overline{X}_{100} - \mu| < k \cdot (\sigma/\sqrt{n})\Big) \geq 1 - (1/k)^2.$$

Ponendo $k \cdot (\sigma/\sqrt{n}) = 0,5$ si ha $k = 0,5 \cdot \sqrt{n/\sigma^2} = 0,5 \cdot \sqrt{100/1,2}$, da cui $k^2 = 0,5^2 \cdot 100/1,2 = 20,83$. Pertanto

$$\mathsf{P}(|\overline{X}_{100} - \mu| < 0,5) \geq 1 - (1/k)^2 = 1 - (1/20,83) = 0,952.$$

c) Si vuole trovare n tale per cui

$$\mathsf{P}(|\overline{X}_n - \mu| < 0,2) \geq 0,95.$$

Dalla disuguaglianza di Tcebycheff si ha:

$$\mathsf{P}(|\overline{X}_n - \mu| < k \cdot (\sigma/\sqrt{n})) \geq 1 - (1/k)^2,$$

da cui si ottiene $1 - (1/k)^2 = 0,95$, e dunque $k^2 = 20$. Pertanto, considerando che

$$0,2 = k\frac{\sigma}{\sqrt{n}}\ , \quad \text{si ottiene} \quad n = \frac{k^2\sigma^2}{(0,2)^2} = \frac{20 \cdot 1,2}{0,04} = 600.$$

Esercizio D. Indicando con n la numerosità del campione da esaminare, con $\overline{X}_n$ la media campionaria delle durate e con μ la media effettiva del nuovo processo, si vuole determinare n tale per cui

$$\mathsf{P}\Big(|\overline{X}_n - \mu| < 50\Big) \geq 0,99.$$

Applicando la disuguaglianza di Tchebycheff alla variabile aleatoria $\overline{X}_n$, si può scrivere:

$$\mathsf{P}\Big(|\overline{X}_n - \mathrm{E}(\overline{X}_n)| < k\sqrt{\mathrm{Var}(\overline{X}_n)}\Big) \geq 1 - \frac{1}{k^2} = 0,99,$$

dove $\mathrm{E}(\overline{X}_n) = \mu$ e $\mathrm{Var}(\overline{X}_n) = 250^2/n$. Per cui, $k = 10$ e

$$50 = k\sqrt{\mathrm{Var}(\overline{X}_n)} = k\sqrt{250^2/n}, \quad \text{da cui} \quad n = 2500.$$