Universidade de Brasília
Departamento de Matemática

Simetrias e Médias de Quocientes de Somas Parciais

L. Cioletti

22 de Julho de 2025

Resumo

Dada uma sequência de variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$ independentes e identicamente distribuídas (iid), investigamos o problema de calcular o valor esperado do quociente entre as somas parciais, S_m e S_n, com $m \leqslant n$. Sob certas hipóteses de integrabilidade, mostramos que

$$\mathbb{E}\left[\frac{S_m}{S_n}\mathbb{1}_{\{S_n\neq 0\}}\right] = \frac{m}{n}\cdot\mathbb{P}(S_n \neq 0).$$

Como veremos a conclusão emerge de um clássico argumento de simetria que, módulo a condição de integrabilidade, independe da escolha da distribuição de X_n. Nosso objetivo é apresentar a prova deste fato em detalhes e estabelecer as condições de integrabilidade para a validade da fórmula. Em particular, vamos mostrar que esta condição de integrabilidade é sempre satisfeita nos casos em que as v.a.'s possuem sinal definido, independentemente de terem esperança finita ou não.

O texto traz vários exemplos, alguns onde a igualdade acima ocorre em casos onde $\mathbb{P}(S_n \neq 0) < 1$. Outros onde $\mathbb{P}(S_n \neq 0) = 1$ e também apresentamos um contra-exemplo usando v.a.'s uniformemente distribuídas em $(-1, 1)$, onde temos $\mathbb{P}(S_n \neq 0) = 1$, mas a fórmula acima não é válida; evidenciando a importância da condição de integrabilidade.

Ressaltamos que este texto consiste, basicamente, em uma discussão detalhada de um exercício simples, comumente encontrado nos livros introdutórios de Teoria de Probabilidade em nível de graduação. O conteúdo é de caráter exclusivamente expositivo e didático. As ideias apresentadas são baseadas em resultados clássicos e muito bem-conhecidos. Assim, não há aqui nenhuma contribuição original ou inovadora.

Palavras-chave: Quocientes de somas parciais, variáveis aleatórias independentes, distribuições invariantes por permutações, Teoria da Probabilidade.

1 Introdução

O estudo de somas de variáveis aleatórias independentes e identicamente distribuídas (iid) está entre os mais importantes dentro da Teoria de Probabilidade, com resultados fundamentais como a Lei Forte dos Grandes Números e o Teorema do Limite Central, que descrevem o comportamento assintótico de somas parciais, devidamente normalizadas, veja referência [1].

Neste artigo estudamos uma questão relacionada: o valor esperado do quociente de somas parciais, $\mathbb{E}[S_m/S_n]$, onde $S_n \equiv X_1 + \ldots + X_n$ e $\{X_n\}_{n\in\mathbb{N}}$ é uma sequência de v.a.'s idd.

Em geral, o cálculo do valor esperado de quocientes de varáveis aleatórias é bastante delicado. Embora o caso que estamos considerando aqui seja muito simples e a fórmula $\mathbb{E}[S_m/S_n] = m/n$, tenha um apelo intuitivo muito grande, sua validade depende criticamente de como a distribuição de S_n se comporta, quando esta v.a. está próxima de zero. Grosseiramente falando, o resultado principal, afirma que

$$\mathbb{E}\left[\frac{S_m}{S_n}\mathbb{1}_{\{S_n\neq 0\}}\right] = \frac{m}{n}\cdot\mathbb{P}(S_n\neq 0), \qquad 1\leqslant m\leqslant n. \tag{1}$$

desde que certas condições de integrabilidade sejam satisfeitas. É crucial notar que esta condição é imposta sobre o próprio quociente, e não necessariamente sobre as variáveis X_j. De fato, como mostraremos no Teorema 7, envolvendo a distribuição de Cauchy, o resultado principal pode ser válido mesmo quando as variáveis X_j não possuem esperança finita.

2 Motivações, Exemplos e Contra-Exemplos

Na seção anterior, comentamos brevemente qual é a natureza do resultado central que pretendemos provar ao longo do texto. Antes de prosseguirmos para os enunciados formais e para as demonstrações, vamos apresentar nesta seção alguns exemplos que vão fornecer a intuição e o motivo por trás de cada uma das hipóteses que serão consideradas.

2.1 Quocientes de Somas Parciais de Bernoullis

Sejam X_1, X_2 v.a.'s independentes com distribuição de Bernoulli de parâmetro $p\in(0,1)$, isto é, $\mathbb{P}(X_j=1)=p$ e $\mathbb{P}(X_j=0)=1-p$, para $j=1,2$. Defina

$$Y_{1,2}\equiv\begin{cases}\dfrac{X_1}{X_1+X_2}, & \text{se } X_1+X_2\neq 0;\\[2ex] 0, & \text{se } X_1+X_2=0.\end{cases} \qquad \text{e} \qquad Y_{2,2}\equiv\begin{cases}\dfrac{X_2}{X_1+X_2}, & \text{se } X_1+X_2\neq 0;\\[2ex] 0, & \text{se } X_1+X_2=0.\end{cases}$$

Observe que neste caso é necessário considerar variáveis aleatórias como $Y_{1,2}$ e $Y_{2,2}$ porque a rigor, não é possível olhar para os quocientes

$$\frac{X_1}{X_1+X_2} \qquad \text{e} \qquad \frac{X_2}{X_1+X_2}$$

como variáveis aleatórias.

Por questão de conveniência podemos ver $Y_{1,2}$ como a composição da função $g : \mathbb{R}^2 \to \mathbb{R}$ com o vetor aleatório $X \equiv (X_1, X_2)$, onde

$$g(x_1, x_2) \equiv \begin{cases} \dfrac{x_1}{x_1 + x_2}, & \text{se} \quad x_1 + x_2 \neq 0; \\ \\ 0, & \text{caso contrário.} \end{cases} \tag{2}$$

Assim segue das propriedades elementares da esperança que

$$\begin{aligned}
\mathbb{E}[Y_{1,2}] &= \mathbb{E}[g(X_1, X_2)] \\
&= \sum_{k_1=0}^{1} \sum_{k_2=0}^{1} g(k_1, k_2) \mathbb{P}\big(\{X_1 = k_1\} \cap \{X_2 = k_2\}\big). \\
&= \sum_{k_1=0}^{1} \sum_{k_2=0}^{1} g(k_1, k_2) \mathbb{P}(X_1 = k_1)\, \mathbb{P}(X_2 = k_2) \\
&= g(0,0)(1-p)^2 + g(0,1)(1-p)p + g(1,0)p(1-p) + g(1,1)p^2. \\
&= p(1-p) + \frac{p^2}{2} \;=\; p - \frac{p^2}{2} \;=\; p\left(1 - \frac{p}{2}\right).
\end{aligned} \tag{3}$$

Observe que o valor médio calculado acima satisfaz $p(1 - p/2) < 1/2$, já que estamos assumindo que $p \in (0, 1)$.

Outra observação importante é que, neste caso, temos

$$\begin{aligned}
\mathbb{P}\big(\{S_2 = 0\}\big) = \mathbb{P}\big(\{X_1 + X_2 = 0\}\big) &= \mathbb{P}\big(\{X_1 = 0\} \cap \{X_2 = 0\}\big) \\
&= \mathbb{P}\big(\{X_1 = 0\}\big) \cdot \mathbb{P}\big(\{X_2 = 0\}\big) \\
&= (1-p)^2.
\end{aligned}$$

Consequentemente,

$$\frac{1}{2} \cdot \mathbb{P}(\{S_2 \neq 0\}) = \frac{1 - (1-p)^2}{2} = \frac{2p - p^2}{2} = p - \frac{p^2}{2} = p\left(1 - \frac{p}{2}\right).$$

Logo, segue da igualdade acima e de (3) que

$$\mathbb{E}[Y_{1,2}] = \frac{1}{2} \cdot \mathbb{P}(\{S_2 \neq 0\}).$$

Vale a pena observar também que os vetores (X_1, X_2) e (X_2, X_1) possuem a mesma função de distribuição e portanto, temos

$$\mathbb{E}[Y_{2,2}] = \mathbb{E}[g(X_2, X_1)] = \mathbb{E}[g(X_1, X_2)] = \frac{1}{2} \cdot \mathbb{P}(\{S_2 \neq 0\}).$$

2.2 Quocientes de Somas Parciais de Uniformes em $(0,1)$

Nesta seção vamos considerar um exemplo semelhante ao da subseção anterior, só que agora com as variáveis aleatórias sendo absolutamente contínuas. Mais precisamente, vamos considerar o caso em que X_1 e X_2 são v.a.'s independentes e com distribuição uniforme no intervalo aberto $(0,1)$. Ou seja

$$X_1 \sim U(0,1) \qquad \text{e} \qquad X_2 \sim U(0,1).$$

Neste caso, é claro que $\mathbb{P}(\{X_1 + X_2 = 0\}) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0) = 0$. Além do mais, segue diretamente das definições de $Y_{1,2}$ e $Y_{2,2}$ que

$$Y_{1,2} = \frac{X_1}{X_1 + X_2} \qquad \text{e} \qquad Y_{2,2} = \frac{X_2}{X_1 + X_2}, \qquad \text{quase certamente.}$$

No que segue, mostramos que

$$\mathbb{E}\left[\frac{X_1}{X_1 + X_2}\right] = \frac{1}{2} = \mathbb{E}\left[\frac{X_2}{X_1 + X_2}\right].$$

Primeiro, vamos verificar que a condição de integrabilidade $\mathbb{E}[|Y_{1,2}|] < +\infty$ é satisfeita. De fato, já que $0 < X_1, X_2 < 1$, e que (X_1, X_2) possui função de densidade conjunta dada por $f(x,y) = \mathbb{1}_{(0,1)\times(0,1)}(x,y)$, então

$$\begin{aligned}
\mathbb{E}\left[\left|\frac{X_1}{X_1 + X_2}\right|\right] = \mathbb{E}\left[\frac{X_1}{X_1 + X_2}\right] &= \mathbb{E}[g(X_1, X_2)] \\
&= \int_{\mathbb{R}^2} g(x,y) \cdot \mathbb{1}_{(0,1)\times(0,1)}(x,y)\, dx\, dy \\
&= \int_0^1 \int_0^1 \frac{x}{x+y}\, dx\, dy \\
&\leqslant \int_0^1 \int_0^1 \frac{x}{x}\, dx\, dy \\
&= 1,
\end{aligned}$$

onde a função g que aparece na segunda igualdade acima é a função definida em (2). De maneira completamente análoga, podemos verificar que $Y_{2,2}$ é integrável.

Para finalizar o cálculo basta observar que

$$\frac{X_1}{X_1 + X_2} + \frac{X_2}{X_1 + X_2} = 1,$$

tomar a esperança em ambos lados da igualdade acima e usar que cada uma das parcelas que aparecem no lado esquerdo da igualdade acima possuem esperança finita e que elas coincidem, já que

$$\mathbb{E}\left[\frac{X_1}{X_1 + X_2}\right] = \mathbb{E}[g(X_1, X_2)] = \mathbb{E}[g(X_2, X_1)] = \mathbb{E}\left[\frac{X_2}{X_1 + X_2}\right]$$

e consequentemente,

$$\mathbb{E}\left[\frac{X_1}{X_1 + X_2}\right] = \frac{1}{2} = \mathbb{E}\left[\frac{X_2}{X_1 + X_2}\right].$$

2.3 Uniformes Simétricas e a Condição de Integrabilidade

Nesta seção vamos considerar um exemplo semelhante ao da subseção anterior, mas agora vamos supor que

$$X_1 \sim U(-1,1) \qquad \text{e} \qquad X_2 \sim U(-1,1).$$

Neste caso, também é possível verificar que tem probabilidade zero do vetor aleatório (X_1, X_2) tomar valores na reta $H \equiv \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 0\}$. Isto é,

$$\mathbb{P}(\{X_1 + X_2 = 0\}) \,=\, 0.$$

De fato, o vetor aleatório (X_1, X_2) tem função de densidade conjunta dada por

$$f(x,y) = \frac{1}{4} \cdot \mathbb{1}_{(-1,1)\times(-1,1)}(x,y), \qquad \forall (x,y) \in \mathbb{R}^2.$$

Portanto,

$$\begin{aligned}
\mathbb{P}(\{X_1 + X_2 = 0\}) = \mathbb{P}((X_1, X_2) \in H) &= \int_H f(x,y)\, dx\, dy \\
&= \frac{1}{4} \int_{(-1,1)\times(-1,1)\cap H} dx\, dy = 0,
\end{aligned} \tag{4}$$

onde na última igualdade usamos que qualquer boreliano contido em uma reta em $\mathbb{R}^2$ tem medida de Lebesgue (bi-dimensional) nula.

Como consequência da igualdade estabelecida acima podemos afirmar que as duas parcelas que aparecem no lado esquerdo da próxima igualdade estão bem-definidas, quase certamente, e vale a seguinte identidade

$$\frac{X_1}{X_1 + X_2} + \frac{X_2}{X_1 + X_2} = \mathbb{1}_{\{(X_1, X_2) \notin H\}}.$$

Neste caso, embora siga da igualdade acima e de (4) que

$$\mathbb{E}\left[\frac{X_1}{X_1 + X_2} + \frac{X_2}{X_1 + X_2}\right] = \mathbb{E}\left[\mathbb{1}_{\{(X_1, X_2) \notin H\}}\right] = \mathbb{P}((X_1, X_2) \notin H) = 1,$$

não é possível prosseguir no cálculo decompondo a esperança, no lado esquerdo da igualdade acima, como soma de duas esperanças porque cada uma das parcelas que aparecem nesta esperança, não possuem esperança finita. Para verificar que esta última afirmação é verdadeira, basta mostrar que

$$\mathbb{E}[|g(X_1, X_2)|] = \int_{\mathbb{R}^2} |g(x,y)| \cdot f(x,y)\, dx\, dy = +\infty,$$

onde g é novamente a função definida em (2) e f é a função de densidade conjunta do vetor aleatório (X_1, X_2).

Note que mostrar que a esperança acima diverge é equivalente, pelo Teorema de Tonelli, a mostrar que

$$\int_{-1}^{1} \int_{-1}^{1} \frac{|x|}{|x+y|}\, dx\, dy = \int_{-1}^{1} |x| \left[\int_{-1}^{1} \frac{1}{|x+y|}\, dy \right] dx = +\infty.$$

Para provar que a integral acima diverge é suficiente mostrar que para cada $x \in (-1,1)$ fixado, temos

$$\int_{-1}^{1} \frac{1}{|x+y|}\, dy = +\infty.$$

Para fazer este cálculo, podemos usar o teorema de mudança de variáveis para integrais de Lebesgue de funções não-negativas e o Teorema da Convergência Monótona como segue

$$\begin{aligned}
\int_{-1}^{1} \frac{1}{|x+y|}\, dy = \int_{-1+x}^{1+x} \frac{1}{|t|}\, dt &= \lim_{n\to\infty} \int_{\mathbb{R}} \left[\mathbb{1}_{[-1+x,-\frac{1}{n}]}(t)\frac{1}{|t|} + \mathbb{1}_{[\frac{1}{n},1+x]}(t)\frac{1}{|t|} \right] dt. \\
&= \lim_{n\to\infty} \left[\int_{-1+x}^{-\frac{1}{n}} \frac{1}{-t}\, dt + \int_{\frac{1}{n}}^{1+x} \frac{1}{t}\, dt \right] \\
&= \lim_{n\to\infty} \left[-\int_{-1+x}^{-\frac{1}{n}} \frac{1}{t}\, dt + \int_{\frac{1}{n}}^{1+x} \frac{1}{t}\, dt \right] \\
&= \lim_{n\to\infty} \left[-\ln|t| \Big|_{-1+x}^{-\frac{1}{n}} + \ln(t)\Big|_{\frac{1}{n}}^{1+x} \right] \\
&= \lim_{n\to\infty} \left[-\ln\left|-\frac{1}{n}\right| + \ln|-1+x| + \ln(1+x) - \ln\left(\frac{1}{n}\right) \right] \\
&= \ln|-1+x| + \ln(1+x) - 2\lim_{n\to\infty} \left[\ln\left(\frac{1}{n}\right) \right] \\
&= \ln|-1+x| + \ln(1+x) + 2\lim_{n\to\infty} \ln(n) \\
&= +\infty.
\end{aligned}$$

3 Provas dos Resultados Principais

Proposição 1. Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sequência de v.a.'s, independentes e identicamente distribuídas. Fixado $k \in \mathbb{N}$ e π uma permutação arbitrária do conjunto $\{1,\ldots,k\}$, temos que $(X_1,\ldots,X_k) \overset{d}{=} (X_{\pi(1)},\ldots,X_{\pi(k)})$.

Prova. Considere os seguintes vetores aleatórios $X \equiv (X_1,\ldots,X_k)$ e $Y \equiv (X_{\pi(1)},\ldots,X_{\pi(k)})$. Para mostrar que eles possuem a mesma função de distribuição é suficiente pelo Teorema de Lévy, observar que eles possuem a mesma função característica. De fato, para qualquer vetor $t \equiv (t_1,\ldots,t_k) \in \mathbb{R}^k$ temos as seguintes igualdades

$$\phi_X(t) = \mathbb{E}[\exp(i\langle t, X\rangle)] = \mathbb{E}\left[\exp\left(i\sum_{j=1}^{k} t_j X_j\right)\right] = \prod_{j=1}^{k} \mathbb{E}\left[\exp\left(it_jX_j\right)\right]$$

$$= \prod_{j=1}^{k} \mathbb{E}\left[\exp\left(it_jX_{\pi(j)}\right)\right] = \mathbb{E}\left[\exp\left(i\sum_{j=1}^{k} t_j X_{\pi(j)}\right)\right] = \mathbb{E}[\exp(i\langle t, Y\rangle)] = \phi_Y(t). \quad \blacksquare$$

Observação. É claro que também podemos provar a Teorema 1 usando diretamente as propriedades elementares de F_X, função de distribuição do vetor aleatório $X \equiv (X_1, \ldots, X_k)$. Basta notar que esta distribuição se fatora, reordenar os fatores e usar que todas as X_i's são variáveis aleatórias possuindo a mesma distribuição.

Proposição 2. Sejam $\psi : \mathbb{R}^k \to \mathbb{R}$ uma função $\mathscr{B}(\mathbb{R}^k)$-mensurável, $X \equiv (X_1, \ldots, X_k)$ um vetor aleatório com coordenadas iid, π uma permutação arbitrária do conjunto de índices $\{1, \ldots, k\}$ e $Y \equiv (X_{\pi(1)}, \ldots, X_{\pi(k)})$.

- Se $\psi(x) \geqslant 0$, para todo $x \in \mathbb{R}^k$, então $\mathbb{E}[\psi(X)] = \mathbb{E}[\psi(Y)]$.
- Se $\mathbb{E}[|\psi(X)|] < +\infty$, então $\mathbb{E}[\psi(X)] = \mathbb{E}[\psi(Y)]$.

Prova. Seja $F_X : \mathbb{R}^k \to [0, 1]$ a função de distribuição do vetor $X \equiv (X_1, \ldots, X_k)$. Pela Teorema 1 segue que a função de distribuição $F_Y = F_X$. Sob qualquer uma das duas condições $\psi \geqslant 0$ ou $\mathbb{E}[|\psi(X)|] < +\infty$, podemos aplicar a Fórmula de Mudança de Variáveis, veja [1, Theorem 16.13, p.229] e concluir que

$$\mathbb{E}[\psi(X)] = \int_{\mathbb{R}^k} \psi \, d\mu_X = \int_{\mathbb{R}^k} \psi \, d\mu_Y = \mathbb{E}[\psi(Y)],$$

onde $\mu_X, \mu_Y : \mathscr{B}(\mathbb{R}^k) \to [0, 1]$ são as medidas de Lebesgue-Stietjes, sobre os borelianos de $\mathbb{R}^k$, induzidas pelas funções de distribuição F_X e F_Y, respectivamente. ■

Para a prova do próximo resultado vamos precisar de um fato bem-conhecido na Teoria de Probabilidades, cujo enunciado é apresentado abaixo.

Proposição 3. Sejam X e Y variáveis aleatórias. Então

$$X \stackrel{d}{=} Y \qquad \iff \qquad \mathbb{E}[\varphi(X)] = \mathbb{E}[\varphi(Y)], \quad \forall \varphi \in C_b(\mathbb{R}, \mathbb{R}),$$

onde $C_b(\mathbb{R}, \mathbb{R}) \equiv \{f : \mathbb{R} \to \mathbb{R} \ : \ f$ é contínua e $\|f\|_\infty \equiv \sup_{x \in \mathbb{R}} |f(x)| < +\infty\}$, ou seja, o espaço das funções, definidas em toda reta, tomando valores reais, contínuas e limitadas.

Prova. Se $X \stackrel{d}{=} Y$, então segue do Teorema da Mudança de variáveis que $\mathbb{E}[\varphi(X)] = \mathbb{E}[\varphi(Y)]$, para toda função $\varphi \in C_b(\mathbb{R}, \mathbb{R})$. Para a recíproca, basta observar que para cada $t \in \mathbb{R}$ fixado, existe uma sequência de funções contínuas e uniformemente limitadas $\varphi_n : \mathbb{R} \to \mathbb{R}$ tal que $\varphi_n(x) \to \mathbb{1}_{(-\infty, t]}(x)$, quando $n \to \infty$ para todo $x \in \mathbb{R}$. Aplicando o Teorema da Convergência Dominada temos

$$\begin{aligned}\mathbb{P}(X \leqslant t) = \mathbb{E}[\mathbb{1}_{(-\infty,t]}(X)] &= \lim_{n\to\infty} \mathbb{E}[\varphi_n(X)] \\ &= \lim_{n\to\infty} \mathbb{E}[\varphi_n(Y)] = \mathbb{E}[\mathbb{1}_{(-\infty,t]}(Y)] = \mathbb{P}(Y \leqslant t).\end{aligned}$$

Como t pode ser escolhido arbitrariamente em $\mathbb{R}$, podemos concluir que $X \stackrel{d}{=} Y$. ■

Corolário 4. Sejam $\{X_n\}_{n\in\mathbb{N}}$ uma sequência de v.a.'s iid e $g:\mathbb{R}^k\to\mathbb{R}$ a função dada por

$$g(x_1,\ldots,x_k)\equiv\begin{cases}\dfrac{x_1}{x_1+\ldots+x_k}, & \text{se} \quad x_1+\ldots+x_k\neq 0;\\[2ex] 0, & \text{caso contrário.}\end{cases} \tag{5}$$

Para qualquer permutação π do conjunto de índices $\{1,\ldots,k\}$ fixada, temos que as expressões $Y\equiv g(X_1,\ldots,X_k)$ e $Z\equiv g(X_{\pi(1)},\ldots,X_{\pi(k)})$ definem variáveis aleatórias com a mesma distribuição. Em particular, se $\mathbb{P}(\{S_k\neq 0\})=1$, então os seguintes quocientes são variáveis aleatórias bem-definidas, quase certamente, e valem as seguintes igualdades em distribuição

$$\frac{X_i}{S_k}\overset{d}{=}\frac{X_j}{S_k}, \qquad 1\leqslant i,j\leqslant k.$$

Prova. Fixe uma permutação π do conjunto de índices $\{1,\ldots,k\}$ e em seguida, considere as expressões $Y\equiv g(X_1,\ldots,X_k)$ e $Z\equiv g(X_{\pi(1)},\ldots,X_{\pi(k)})$. Para verificar que Y e Z definem variáveis aleatórias, é suficiente mostrar que a função $g:\mathbb{R}^k\to\mathbb{R}$, definida em (5), é uma função $\mathscr{B}(\mathbb{R}^k)$-mensurável. De fato, seja $H\subset\mathbb{R}^k$ o hiperplano determinado pela equação $x_1+\ldots+x_k=0$. Como todo hiperplano de $\mathbb{R}^k$ é um fechado de $\mathbb{R}^k$, então $H\in\mathscr{B}(\mathbb{R}^k)$. É claro que a função g, restrita à $\mathbb{R}^k\setminus H$ é uma função contínua. Logo, segue destas duas observações que g é uma função real $\mathscr{B}(\mathbb{R}^k)$-mensurável.

Para mostrar que $X\overset{d}{=}Y$ vamos usar a Teorema 3, ou seja, vamos mostrar que $\mathbb{E}[\varphi(Y)]=\mathbb{E}[\varphi(Z)]$, para toda $\varphi\in C_b(\mathbb{R},\mathbb{R})$.

Primeiro, observamos que se $\varphi\in C_b(\mathbb{R},\mathbb{R})$, então temos

$$\mathbb{E}[|(\varphi\circ g)(Y)|]\leqslant\mathbb{E}[\|\varphi\|_\infty]=\|\varphi\|_\infty<+\infty.$$

Analogamente, temos $\mathbb{E}[|(\varphi\circ g)(Z)|]<+\infty$. Aplicando a Teorema 2, com $\psi=\varphi\circ g$, podemos concluir que

$$\begin{aligned}\mathbb{E}[\varphi(Y)]=\mathbb{E}[\varphi(g(X_1,\ldots,X_k))]&=\mathbb{E}[(\varphi\circ g)(X_1,\ldots,X_k)]\\&=\mathbb{E}[(\varphi\circ g)(X_{\pi(1)},\ldots,X_{\pi(k)})]\\&=\mathbb{E}[\varphi(g(X_{\pi(1)},\ldots,X_{\pi(k)}))]=\mathbb{E}[\varphi(Z)].\end{aligned}$$

Como a igualdade acima é válida para qualquer função $\varphi\in C_b(\mathbb{R},\mathbb{R})$ segue diretamente da Teorema 3 que Y e Z possuem a mesma distribuição.

Para finalizar, suponha que $\mathbb{P}(S_k\neq 0)=1$. Neste caso o quociente X_j/S_k, com $1\leqslant j\leqslant k$, define uma variável aleatória, quase certamente. Fixe um par de índices $i,j\in\{1,\ldots,k\}$ e considere as permutações π e σ, definidas como segue:

- $\pi(1)=i$, $\pi(i)=1$ e $\pi(r)=r$, para todo $r\in\{1,\ldots,k\}\setminus\{1,i\}$;
- $\sigma(1)=j$, $\sigma(j)=1$ e $\sigma(r)=r$, para todo $r\in\{1,\ldots,k\}\setminus\{1,j\}$;

Então segue da definição da função g, das permutações π e σ e da primeira parte que

$$\frac{X_i}{S_k}=g(X_{\pi(1)},\ldots,X_{\pi(k)})\overset{d}{=}g(X_1,\ldots,X_k)\overset{d}{=}g(X_{\sigma(1)},\ldots,X_{\sigma(k)})=\frac{X_j}{S_k}. \qquad \blacksquare$$

Proposição 5. Se $X \equiv (X_1, \ldots, X_k)$ é um vetor aleatório com coordenadas iid, e a seguinte condição de integrabilidade $\mathbb{E}[|g_{1,k}(X)|] < +\infty$ é satisfeita, então temos

$$\mathbb{E}\left[\sum_{i=1}^{j} g_{i,k}(X)\right] \;=\; \frac{j}{k} \cdot \mathbb{P}(X \notin H), \qquad \forall j \in \{1, \ldots, k\}.$$

Prova. Fixe $j, k \in \mathbb{N}$ com $j \leqslant k$ e considere a função $g_{j,k} : \mathbb{R}^k \to \mathbb{R}$ dada por

$$g_{j,k}(x_1, \ldots, x_k) \equiv \begin{cases} \dfrac{x_j}{x_1 + \ldots + x_k}, & \text{se} \quad x_1 + \ldots + x_k \neq 0; \\[2ex] 0, & \text{caso contrário.} \end{cases}$$

Argumentando de maneira análoga a do Teorema 4, podemos verificar que $g_{j,k}$ é uma função $\mathscr{B}(\mathbb{R}^k)$-mensurável. Observe que segue diretamente da definição de $g_{j,k}$ que

$$\mathbb{1}_{\{X_1+\ldots+X_k \neq 0\}} \;=\; g_{1,k}(X) + \ldots + g_{k,k}(X). \tag{6}$$

Considere a permutação π de $\{1, \ldots, k\}$ dada por $\pi(1) = j$, $\pi(j) = 1$ e $\pi(s) = s$, para todo $s \in \{1, \ldots, k\} \setminus \{1, j\}$. Isto é, a permutação π, troca os índices 1 e j, mas mantém todos os demais pontos fixos. Considere o vetor aleatório $Y \equiv (X_{\pi(1)}, \ldots, X_{\pi(j)}, \ldots, X_{\pi(k)}) = (X_j, \ldots, X_1, \ldots, X_k)$.

Como estamos assumindo que $\mathbb{E}[|g_{1,k}(X)|] < +\infty$, podemos aplicar a Teorema 2 para garantir que valem as seguintes igualdades:

$$\mathbb{E}\big[|g_{1,k}(X)|\big] = \mathbb{E}\big[|g_{1,k}(Y)|\big] \qquad \text{e} \qquad \mathbb{E}[g_{1,k}(X)] = \mathbb{E}[g_{1,k}(Y)]. \tag{7}$$

Note que segue diretamente das definições: do vetor Y, da função $g_{j,k}$ e da última igualdade acima que

$$\mathbb{E}[g_{1,k}(X)] = \mathbb{E}[g_{j,k}(X)], \qquad 1 \leqslant j \leqslant k. \tag{8}$$

Usando a equação (6), a condição de integrabilidade $\mathbb{E}[|g_{1,k}(X)|] < +\infty$, as propriedades elementares da esperança e a identidade acima obtemos

$$\mathbb{P}(X \notin H) = \mathbb{E}\big[\mathbb{1}_{\{X_1+\ldots+X_k \neq 0\}}\big] = \mathbb{E}\big[g_{1,k}(X) + \ldots + g_{k,k}(X)\big] = k\mathbb{E}[g_{1,k}(X)].$$

Usando novamente (8) e a identidade acima, temos finalmente que

$$\mathbb{E}\left[\sum_{i=1}^{j} g_{i,k}(X)\right] = \frac{j}{k} \cdot \mathbb{P}(X \notin H), \qquad 1 \leqslant j \leqslant k. \qquad \blacksquare$$

A seguir sintetizamos em forma de teorema as partes mais importantes da Teorema 5, e também traduzindo os conceitos que envolvem a notação mais abstrata de $g_{i,k}(X)$, para a notação mais intuitiva de S_j/S_k.

Teorema 6. Sejam $\{X_n\}_{n\in\mathbb{N}}$ uma sequência de v.a.'s idd, $k \in \mathbb{N}$, $X \equiv (X_1, \ldots, X_k)$ e $g : \mathbb{R}^k \to \mathbb{R}$ a função definida por (5). Se $\mathbb{E}[|g(X)|] < +\infty$, então temos a seguinte igualdade

$$\mathbb{E}\left[\frac{S_j}{S_k}\mathbb{1}_{\{S_k\neq 0\}}\right] = \frac{j}{k}\cdot \mathbb{P}(S_k \neq 0), \qquad 1 \leqslant j \leqslant k.$$

Em particular, se $\mathbb{P}(S_k \neq 0) = 1$, então a condição de integrabilidade $\mathbb{E}[|g(X)|] < +\infty$, é equivalente à $\mathbb{E}[|X_1/S_k|] < +\infty$ e neste caso temos

$$\mathbb{E}\left[\frac{S_j}{S_k}\right] = \frac{j}{k}, \qquad 1 \leqslant j \leqslant k.$$

Prova. Fixe índices $1 \leqslant j \leqslant k$ e seja H o hiperplano definido na Teorema 5. Se X denota o vetor aleatório $(X_1, \ldots, X_k)$, então temos claramente que $\{X \notin H\} = \{S_k \neq 0\}$. Além do mais, se interpretamos a expressão do lado esquerdo da igualdade abaixo como sendo zero quando $S_k = 0$, então temos:

$$\frac{S_j}{S_k}\mathbb{1}_{\{S_k\neq 0\}} = \sum_{i=1}^{j} g_{i,k}(X).$$

Logo, segue das observações acima e da conclusão da Teorema 5 que

$$\mathbb{E}\left[\frac{S_j}{S_k}\mathbb{1}_{\{S_k\neq 0\}}\right] = \mathbb{E}\left[\sum_{i=1}^{j} g_{i,k}(X)\right] = \frac{j}{k}\cdot \mathbb{P}(X \notin H) = \frac{j}{k}\cdot \mathbb{P}(S_k \neq 0). \tag{9}$$

No caso em que $\mathbb{P}(S_k \neq 0) = 1$, temos que $g(X) = X_1/S_k$, quase certamente e portanto a condição de integrabilidade de $g(X)$ pode ser reescrita como $E[|X_1/S_k|] < +\infty$. Neste caso, segue diretamente de (9) que $\mathbb{E}[S_j/S_k] = j/k$. ■

Exemplo 7 (O caso de v.a.'s não-integráveis). Seja $\{Y_n\}_{n\in\mathbb{N}}$ uma sequência iid com $Y_n \sim \text{Cauchy}(\lambda)$, onde $\lambda > 0$. Ou seja, para cada $n \in \mathbb{N}$, a variável aleatória Y_n possui função de densidade dada por

$$f_{Y_n}(x) = \frac{1}{\pi}\frac{\lambda}{\lambda^2 + x^2}, \qquad \forall x \in \mathbb{R}.$$

Para cada $n \in \mathbb{N}$, considere a v.a. $X_n \equiv |Y_n|$. Como X_n é uma variável aleatória não-negativa, sua esperança está bem-definida. Mas neste caso, podemos verificar diretamente que $\mathbb{E}[X_n] = +\infty$.

Seja $S_k \equiv X_1 + \ldots + X_k$. Como as v.a.'s X_n são independentes, não-negativas e absolutamente contínuas, temos que

$$\mathbb{P}(S_k = 0) = \mathbb{P}\left(\bigcap_{i=1}^{k}\{X_i = 0\}\right) = \prod_{i=1}^{k}\mathbb{P}(\{X_i = 0\}) = 0. \tag{10}$$

Vamos mostrar que, embora $\mathbb{E}[X_n]$ seja infinita, a condição de integrabilidade exigida na hipótese do Teorema 6 é satisfeita. De fato, como $\mathbb{P}(S_k \neq 0) = 1$, a condição a ser verificada é $\mathbb{E}[|X_1/S_k|] < +\infty$. Na ocorrência do evento $\{S_k \neq 0\}$, cada $X_i \geq 0$, portanto temos que $S_k = X_1 + \cdots + X_k \geq X_1$. Assim,

$$0 \leqslant \frac{X_1}{S_k} \leqslant 1, \qquad \text{quase certamente.}$$

Isso implica que a variável aleatória X_1/S_k é limitada. Portanto, sua esperança é finita:

$$\mathbb{E}\left[\left|\frac{X_1}{S_k}\right|\right] = \mathbb{E}\left[\frac{X_1}{S_k}\right] \leqslant \mathbb{E}[1] = 1 < +\infty.$$

Desta forma todas as hipóteses do Teorema 6 são satisfeitas. Lembrando que $\mathbb{P}(S_k \neq 0) = 1$, concluímos que

$$\mathbb{E}\left[\frac{S_j}{S_k}\right] = \frac{j}{k}, \qquad 1 \leqslant j \leqslant k.$$

Exemplo 8 (Variáveis Aleatórias Não-Negativas)**.** O resultado visto no exemplo anterior, envolvendo o valor absoluto de uma Cauchy, pode ser generalizado para o caso em que as variáveis aleatórias X_n são simplesmente v.a.'s iid não-negativas. Na verdade, a condição de integrabilidade do Teorema 6 é sempre satisfeita para **qualquer** sequência de variáveis aleatórias iid não-negativas.

Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sequência iid tal que $\mathbb{P}(X_n \geqslant 0) = 1$. Como cada $X_i \geqslant 0$, a soma $S_k = X_1 + \ldots + X_k$ é sempre maior ou igual a cada uma de suas parcelas. Em particular,

$$X_1 \mathbb{1}_{\{S_k\neq 0\}} \leqslant S_k \mathbb{1}_{\{S_k\neq 0\}}, \text{ quase certamente} \qquad \Longrightarrow \qquad \frac{X_1}{S_k}\mathbb{1}_{\{S_k\neq 0\}} \leqslant 1, \text{ quase certamente.}$$

Note que a última desigualdade acima, implica na validade da hipótese de integrabilidade do Teorema 6. De fato, se X denota o vetor aleatório $(X_1, \ldots, X_k)$ temos da definição da função g dada em (5) e da desigualdade acima que

$$|g(X)| \leqslant \frac{X_1}{S_k}\mathbb{1}_{\{S_k\neq 0\}} \leqslant 1, \qquad \text{quase certamente.}$$

Logo a hipótese do Teorema 6 é satisfeita, independentemente de $\mathbb{E}[X_1]$ ser finita ou não. Portanto, para **qualquer** sequência de v.a.'s iid não-negativas (ou similarmente, não-positivas), temos que:

$$\mathbb{E}\left[\frac{S_j}{S_k}\mathbb{1}_{\{S_k\neq 0\}}\right] = \frac{j}{k}\cdot \mathbb{P}(S_k \neq 0), \qquad 1 \leqslant j \leqslant k.$$

Este resultado revela que a falha na condição de integrabilidade, como a que vimos no exemplo das uniformes simétricas, está intrinsecamente ligada à possibilidade de as variáveis aleatórias não terem sinal definido, quase certamente.

4 Observações Finais e Generalizações

Os principais resultados apresentados nestes texto podem ser imediatamente generalizados para o caso em que a sequência $\{X_n\}_{n\in\mathbb{N}}$ satisfaz as condições de integrabilidade do Teorema 6 e de invariância por permutações (*exchangeable*), no seguinte sentido. Para toda permutação $\pi : \mathbb{N} \to \mathbb{N}$ satisfazendo: existe algum $n_0 \in \mathbb{N}$ tal que $\pi(n) = n$, para todo $n \geqslant n_0$; temos que as sequências $\{X_n\}_{n\in\mathbb{N}}$ e $\{X_{\pi(n)}\}_{n\in\mathbb{N}}$ possuem as mesmas distribuições finito dimensionais. Esta classe de sequências de v.a.'s é um pouco mais geral do que a classe de sequências iid.

Referências

[1] P. Billingsley. *Probability and Measure - Anniversary Edition*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, Anniversary edition, 2012. x+608 pp. ISBN: 978-1-118-12237-2.