Aula 12: Somas de v.a. independentes

Disciplina: PGE950 - Probabilidade

Programa de Pós-graduação em Estatística da UFPE

Prof. Pablo M. Rodriguez

http://www.pablo-rodriguez.org

Universidade Federal de Pernambuco

CCEN Centro de Ciências Exatas e da Natureza

Conteúdo desta aula

- Distribuição da soma de variáveis aleatórias independentes.

Somas de v.a. discretas independentes

Sejam X e Y v. a. independentes, tomando valores em $\mathbb{Z}$ e com f. p. dada por $p_X(i)$ e $p_Y(i)$, respectivamente. A convolução de $p_X(i)$ e $p_Y(i)$ define-se como:

$$(p_X * p_Y)(i) = \sum_{k=-\infty}^{\infty} p_X(k) p_Y(i-k),$$

para $i \in \mathbb{Z}$. Em particular, a função

$$p_X * p_Y$$

é a f.p. da v. a.

$$Z = X + Y.$$

Observação!

Seja $S_n = X_1 + X_2 + \cdots + X_n$ *a soma de* n *v.a. discretas i.i.d. com f.p. comum dada por* $p(x)$. *Note que*

- S_1 *tem f.p. dada por* $p(x)$, *e que*
- $S_n = S_{n-1} + X_n$.

Então, podemos encontrar a $f.p.$ *de* S_n *por indução.*

Exemplo 12.1

Suponha que um dado honesto é lançado duas vezes. Sejam X_1 e X_2 os respectivos resultados de cada lançamento, e seja

$$S_2 = X_1 + X_2.$$

Sabemos que a f.p. marginal de X_1 e de X_2 é dada pela mesma função

i	1	2	3	4	5	6
$p(i)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

$\cdots$ *continuação do Exemplo 12.1.* Então, a f.p. de S_2 é dada pela convolução de $p(i)$ com ela mesma. Isto é,

$$P(S_2 = 2) = p(1)p(1) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36},$$

$$P(S_2 = 3) = p(1)p(2) + p(2)p(1) = \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{2}{36}$$

$$P(S_2 = 4) = p(1)p(3) + p(2)p(2) + p(3)p(1) = \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{3}{36},$$

... e continuando desta forma concluímos que sua f.p. é dada por

i	2	3	4	5	6	7	8	9	10	11	12
$(p * p)(i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Sejam X e Y duas v.a. independentes.

- Se $X \sim B(m, p)$ e $Y \sim B(n, p)$, então
$$X + Y \sim B(m + n, p).$$
- Se $X \sim Poisson(\lambda_1)$ e $Y \sim Poisson(\lambda_2)$, então
$$X + Y \sim Poisson(\lambda_1 + \lambda_2).$$
- Se $X \sim Geometrica(p)$ e $Y \sim Geometrica(p)$, então
$$X + Y \sim BN(p, 2).$$

Somas de v.a. geométricas independentes

Proposição 12.1

Sejam $X_1, \ldots, X_n$ *v.a. independentes tais que*

$$X_i \sim Geom(p_i),$$

para $i \in \{1, 2, \ldots, n\}$. *Suponha que* $p_i \neq p_j$ *para* $i \neq j$, *com* $i, j \in \{1, \ldots, n\}$. *Então, para* $k \geq n$ *temos que*

$$P(S_n = i) = \sum_{k=1}^{n} p_k (1 - p_k)^{i-1} \prod_{j \neq k} \frac{p_j}{p_j - p_k}.$$

Demonstração.

Por indução sobre o valor de $n + i$. Ver em “Probabilidade: um curso moderno com aplicações” de Ross (pág. 316 e 317). □

Somas de v.a. continuas independentes

Sejam X e Y duas v.a. continuas com densidade $f_X(x)$ e $f_Y(y)$, respectivamente. A convolução de f_X e f_Y é dada por

$$(f_X * f_Y)(z) := \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\, dy = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x)\, dx.$$

Observação!

Se X e Y são independentes então $f_X * f_Y$ é a densidade da v.a. $Z = X + Y$.

Teorema 12.1

Sejam X e Y v.a. continuas independentes com densidades f_X e f_Y, resp. Então

$$F_{X+Y}(z) = \int_{-\infty}^{\infty} F_X(z-y) \cdot f_Y(y)\, dy.$$

Demonstração.

$$F_{X+Y}(z) = P(X+Y \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x) \cdot f_Y(y)\, dx\, dy.$$

Então,

$$F_{X+Y}(z) = \int_{-\infty}^{\infty} f_Y(y) \cdot \underbrace{\left(\int_{-\infty}^{z-y} f_X(x)\, dx \right)}_{F_X(z-y)} dy.$$

$\square$

Observação!

Derivando a equação anterior com relação a z *obtemos a densidade de* $X+Y$; *i.e.,* $f_{X+Y}(z)=\int_{-\infty}^{\infty} f_X(z-y)\cdot f_Y(y)\,dy$.

Observação!

Como $P(X+Y\leq z)=\int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_X(x)\cdot f_Y(y)\,dy\,dx$, *temos que a densidade de* $X+Y$ *é* $f_{X+Y}(z)=\int_{-\infty}^{\infty} f_Y(z-x)\cdot f_X(x)\,dx$.

Somas de v.a. uniformes independentes

Exemplo 12.2

Sejam X e Y v.a. i.i.d. com $X \sim U(0,1)$. Como

$$f_X(a) = f_Y(a) = \begin{cases} 1, & 0 < a < 1, \\ 0, & \textit{caso contrário.} \end{cases}$$

temos que

$$f_{X+Y}(a) = \int_0^1 f_X(a-y) \cdot f_Y(y)\, dy = \int_0^1 f_X(a-y)\, dy.$$

Agora,

- *para* $0 \leq a \leq 1$*:* $f_{X+Y}(a) = \int_0^a dy = a,$
- *para* $1 < a < 2$*:* $f_{X+Y}(a) = \int_{a-1}^1 dy = 2 - a.$

$\cdots$ *continuação do Exemplo 12.2.* Portanto

$$f_{X+Y}(a) = \begin{cases} a, & 0 \leq a \leq 1, \\ 2-a, & 1 < a < 2, \\ 0, & \text{caso contrário.} \end{cases}$$

Observação!

Dizemos que a v.a. $X+Y$ tem distribuição triangular.

Proposição 12.2

Sejam $X_1, X_2, \ldots, X_n$ *v.a. i.i.d. com* $X_1 \sim U(0,1)$. *Se*

$$F_n(x) := P(X_1 + \ldots + X_n \leq x)$$

então

$$F_n(x) = \frac{x^n}{n!} \tag{1}$$

para $x \in [0,1]$.

Prova da Proposição 12.2. Vamos provar (1) por indução em n. Seja $x \in [0,1]$. Como

$$F_1(x) = P(X_1 \leq x) = x$$

temos que o resultado é valido para $n = 1$. Suponha que (1) é valida para $n-1$, isto é

$$F_{n-1}(x) = \frac{x^{n-1}}{(n-1)!}$$

Para provar que ela é valida para n, escrevemos

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n-1} X_i + X_n$$

e notamos que para $x \in [0,1]$ temos

$$F_n(x) = \int_0^1 F_{n-1}(x-y) \cdot f_{X_n}(y)\,dy = \int_0^x \frac{(x-y)^{n-1}}{(n-1)!}\,dy = \frac{x^n}{n!}$$

Exemplo 12.3

Sejam $X_1, X_2, \dots$ *v.a. i.i.d. com* $X_1 \sim U(0,1)$. *Seja*

$$N = \min\{n > 1 : X_1 + \cdots + X_n > 1\}.$$

Note que para $n > 0$

$$\{N > n\} = \{X_1 + \cdots X_n \leq 1\}.$$

Então,

$$P(N > n) = P(X_1 + \cdots X_n \leq 1) = F_n(1) = \frac{1}{n!}$$

e

$$P(N = n) = P(N > n-1) - P(N > n) = \frac{n-1}{n!}.$$

Lembrete: distribuição gama

Uma v.a. continua X tem distribuição gama com parâmetros (α, λ), para $\alpha > 0$ e $\lambda > 0$, denotamos $X \sim Gama(\alpha, \lambda)$, se sua densidade é dada por

$$f(x) = \begin{cases} \dfrac{\lambda\, e^{-\lambda x} (\lambda x)^{\alpha - 1}}{\Gamma(\alpha)}, & 0 < x < \infty \\ 0, & \text{caso contrário,} \end{cases}$$

onde

$$\Gamma(\alpha) := \int_0^\infty e^{-y}\, y^{\alpha - 1}\, dy$$

é a função gama.

Lembrete: função gama

- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.
- Se $\alpha \in \mathbb{N}$, então $\Gamma(\alpha) = (\alpha - 1)!$
- $\Gamma(1/2) = \sqrt{\pi}$.
- Se $X \sim Exp(\lambda)$ então $X \sim Gama(1, \lambda)$.

Somas de v.a. gama independentes

Proposição 12.3

Se X e Y são v.a. independentes tais que

$$X \sim Gama(s, \lambda)$$

e

$$Y \sim Gama(t, \lambda),$$

então

$$X + Y \sim Gama(s + t, \lambda).$$

Exemplo 12.4

Sejam $X_1, X_2, \dots X_n$ v.a. i.i.d. com $X_1 \sim Exp(\lambda)$. Então

$$X_1 + \cdots + X_n \sim Gama(n, \lambda).$$

Uma propriedade útil!

Se X é uma v.a. com densidade f_X, então a densidade de $Y = X^2$ pode ser obtida da seguinte forma: para $y \geq 0$

$$F_Y(y) := P(Y \leq y) = P(X^2 \leq y) = P(|X| \leq \sqrt{y})$$

Isto é,

$$F_Y(y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Logo, derivando com relação a y obtemos

$$f_Y(y) = \frac{1}{2\sqrt{y}} \left(f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right).$$

Distribuição chi-quadrado

Dizemos que uma v.a. X tem distribuição chi-quadrado com n graus de liberdade, denotamos $X \sim \mathcal{X}^2(n)$, se sua densidade é dada por

$$f(x) = \begin{cases} \dfrac{e^{-x/2} x^{(n/2)-1}}{2^{n/2}\Gamma(n/2)}, & 0 < x < \infty \\ 0, & \text{caso contrário,} \end{cases}$$

onde $\Gamma(\alpha)$ é a função gama.

Observação!

Note que se $X \sim \mathcal{X}^2(n)$ então $X \sim Gama(n/2, 1/2)$.

Distribuição chi-quadrado como soma de v.a. i.i.d.

Sejam $X_1, X_2, \ldots, X_n$ v.a. i.i.d. tais que $X_1 \sim \mathcal{N}(0,1)$. Seja

$$Y = \sum_{i=1}^{n} X_i^2.$$

Então $Y \sim \mathcal{X}^2(n)$. De fato,

$$f_{X_1^2}(y) = \frac{1}{2\sqrt{y}}[f_{X_1}(\sqrt{y}) + f_{X_1}(-\sqrt{y})]$$

e

$$f_{X_1^2}(y) = \frac{1}{2\sqrt{y}} \cdot \frac{2}{\sqrt{2\pi}} \cdot e^{-y/2} = \frac{(1/2)e^{-y/2}(y/2)^{1/2-1}}{\underbrace{\sqrt{\pi}}_{\Gamma(1/2)}}.$$

Logo $f_{X_1^2}(y) \sim Gama(1/2, 1/2)$, e portanto $Y \sim Gama(n/2, 1/2)$.

Universidade Federal de Pernambuco
CCEN
Centro de Ciências Exatas e da Natureza

Distribuição de Rayleigh

Se X e Y são v.a. i.i.d. com $X \sim \mathcal{N}(0,1)$ então

$$X^2 + Y^2 \sim Gama(1, 1/2).$$

Se $Z = \sqrt{X^2 + Y^2}$ então sua densidade é dada por

$$f_Z(z) = \begin{cases} ze^{-z^2/2}, & z \geq 0 \\ 0, & \text{caso contrário.} \end{cases}$$

e é chamada de distribuição de Rayleigh de parâmetro 1.

Observação!

Se escolhemos um ponto P no plano xy com coordenadas (X, Y) então Z é a distância de P à origem!

Somas de v.a. normais independentes

Proposição 12.4

Se $X_1, \ldots, X_n$ *são v.a. independentes com distribuição* $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, *então*

$$\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right).$$

Ideia para a prova da Proposição 12.4.

- Suponha que $X \sim \mathcal{N}(0, \sigma^2)$ e $Y \sim \mathcal{N}(0, 1)$. Então, prove que

$$X + Y \sim \mathcal{N}(0, 1 + \sigma^2).$$

- Considere agora $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ e $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Então,

$$X_1 + X_2 = \sigma_2 \left[\frac{X_1 - \mu_1}{\sigma_2} + \frac{X_2 - \mu_2}{\sigma_2} \right] + \mu_1 + \mu_2$$

- Como

$$\frac{X_1 - \mu_1}{\sigma_2} \sim \mathcal{N}(0, \sigma_1^2/\sigma_2^2) \quad \text{e} \quad \frac{X_2 - \mu_2}{\sigma_2} \sim \mathcal{N}(0, 1)$$

então

$$\frac{X_1 - \mu_1}{\sigma_2} + \frac{X_2 - \mu_2}{\sigma_2} \sim \mathcal{N}(0, 1 + \sigma_1^2/\sigma_2^2).$$

Concluímos que

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

... continuação da ideia para a prova da Proposição 12.4.

Em geral,

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n-1} X_i + X_n$$

e aplicando indução em n podemos provar que

$$\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right).$$

Bom estudo!