Expressões aproximadas e desigualdades

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 08 de outubro de 2024

Roteiro

Roteiro

Motivação

Nós já observamos que a fim de calcular $\mathbb{E}(Z)$ ou $\text{Var}(Z)$, em que $Z = H(X, Y)$, não necessitamos achar a distribuição de probabilidades de Z, pois poderemos trabalhar diretamente com a distribuição de probabilidades de $(X, Y)^\top$.

Motivação

Se a função H for bem complicada, o cálculo das esperanças e variâncias pode conduzir a integrações (ou somatórios) que são bastante difíceis. Por isso, aproximações são muito úteis.

Roteiro

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$ é uma série de funções expressa da seguinte forma

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$ é uma série de funções expressa da seguinte forma

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \tag{1}$$

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$ é uma série de funções expressa da seguinte forma

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n, \tag{1}$$

em que f é uma função conhecida,

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$ é uma série de funções expressa da seguinte forma

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \tag{1}$$

em que f é uma função conhecida, derivável

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$ é uma série de funções expressa da seguinte forma

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \tag{1}$$

em que f é uma função conhecida, derivável e $f^{(n)}$ é a n-ésima derivada de f.

Série de Taylor

A série de Taylor (1715) em torno do ponto $x = a$ é uma série de funções expressa da seguinte forma

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \tag{1}$$

em que f é uma função conhecida, derivável e $f^{(n)}$ é a n-ésima derivada de f.

Série de Taylor

Observação

Em (1),

Série de Taylor

Observação

Em (1), quando $a = 0$,

Observação

Em (1), quando $a = 0$, a série também é conhecida como série de Maclaurin.

Série de Taylor

Observação

Em (1), quando $a = 0$, a série também é conhecida como série de Maclaurin.

Série de Taylor

Notem que, a série de Taylor (1715)

Série de Taylor

Notem que, a série de Taylor (1715) pode ser reescrita da seguinte forma

Série de Taylor

Notem que, a série de Taylor (1715) pode ser reescrita da seguinte forma

$$f(x) = P_n(x) + R_n(x), \tag{2}$$

Série de Taylor

Notem que, a série de Taylor (1715) pode ser reescrita da seguinte forma

$$f(x) = P_n(x) + R_n(x), \tag{2}$$

em que

Notem que, a série de Taylor (1715) pode ser reescrita da seguinte forma

$$f(x) = P_n(x) + R_n(x), \tag{2}$$

em que

$$P_n(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i \text{ e } R_n(x) = \sum_{i=n+1}^{\infty} \frac{f^{(i)}(a)}{i!}(x-a)^i.$$

Série de Taylor

Notem que, a série de Taylor (1715) pode ser reescrita da seguinte forma

$$f(x) = P_n(x) + R_n(x), \tag{2}$$

em que

$$P_n(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i \text{ e } R_n(x) = \sum_{i=n+1}^{\infty} \frac{f^{(i)}(a)}{i!}(x-a)^i.$$

Série de Taylor

$P_n(x)$ é denominado de polinômio de Taylor de grau n e $R_n(x)$ é o resto da Série de Taylor.

Série de Taylor

$P_n(x)$ é denominado de polinômio de Taylor de grau n e $R_n(x)$ é o resto da Série de Taylor. Taylor também mostrou que

Série de Taylor

$P_n(x)$ é denominado de polinômio de Taylor de grau n e $R_n(x)$ é o resto da Série de Taylor. Taylor também mostrou que

$$R_n(x) = \frac{f^{(n+1)}(\tau)}{(n+1)!}(x-a)^n,$$

Série de Taylor

$P_n(x)$ é denominado de polinômio de Taylor de grau n e $R_n(x)$ é o resto da Série de Taylor. Taylor também mostrou que

$$R_n(x) = \frac{f^{(n+1)}(\tau)}{(n+1)!}(x-a)^n,$$

para algum valor de $a \leq \tau \leq x$.

Série de Taylor

$P_n(x)$ é denominado de polinômio de Taylor de grau n e $R_n(x)$ é o resto da Série de Taylor. Taylor também mostrou que

$$R_n(x) = \frac{f^{(n+1)}(\tau)}{(n+1)!}(x-a)^n,$$

para algum valor de $a \leq \tau \leq x$.

Roteiro

Expressões aproximadas da média e da variância

Seja X uma variável aleatória com $\mathbb{E}(X) = \mu$ e $\mathrm{Var}(X) = \sigma^2$. Suponham que $Y = H(X)$.

Expressões aproximadas da média e da variância

Seja X uma variável aleatória com $\mathbb{E}(X) = \mu$ e $\text{Var}(X) = \sigma^2$. Suponham que $Y = H(X)$. Então,

$$\mathbb{E}(Y) \approx H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 \text{ e } \text{Var}(Y) \approx \left[H^{(1)}(\mu)\right]^2 \sigma^2. \qquad (3)$$

Expressões aproximadas da média e da variância

Seja X uma variável aleatória com $\mathbb{E}(X) = \mu$ e $\text{Var}(X) = \sigma^2$. Suponham que $Y = H(X)$. Então,

$$\mathbb{E}(Y) \approx H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 \text{ e } \text{Var}(Y) \approx \left[H^{(1)}(\mu)\right]^2 \sigma^2. \tag{3}$$

Demonstração 1. Usando (1),

Expressões aproximadas da média e da variância

Demonstração 1. Usando (1), até o segundo termo

Demonstração 1. Usando (1), até o segundo termo e em torno de $X = \mu$,

Expressões aproximadas da média e da variância

Demonstração 1. Usando (1), até o segundo termo e em torno de $X = \mu$,

nós temos que

Expressões aproximadas da média e da variância

Demonstração 1. Usando (1), até o segundo termo e em torno de $X = \mu$, nós temos que

$$
\begin{aligned}
Y = H(X) &= \sum_{n=0}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n \\
&= \frac{H^{(0)}(\mu)}{0!}(X-\mu)^0 + \frac{H^{(1)}(\mu)}{1!}(X-\mu)^1 \\
&+ \frac{H^{(2)}(\mu)}{2!}(X-\mu)^2 + \underbrace{\sum_{n=3}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n}_{R_1}.
\end{aligned}
$$

Expressões aproximadas da média e da variância

Demonstração 1. Usando (1), até o segundo termo e em torno de $X = \mu$, nós temos que

$$
\begin{aligned}
Y = H(X) &= \sum_{n=0}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n \\
&= \frac{H^{(0)}(\mu)}{0!}(X-\mu)^0 + \frac{H^{(1)}(\mu)}{1!}(X-\mu)^1 \\
&+ \frac{H^{(2)}(\mu)}{2!}(X-\mu)^2 + \underbrace{\sum_{n=3}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n}_{R_1}.
\end{aligned}
$$

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$Y = H(\mu) + H^{(1)}(\mu)(X-\mu) + \frac{H^{(2)}(\mu)}{2}(X-\mu)^2 + R_1.$$

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$Y = H(\mu) + H^{(1)}(\mu)(X - \mu) + \frac{H^{(2)}(\mu)}{2}(X - \mu)^2 + R_1.$$

Aplicando a esperança,

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$Y = H(\mu) + H^{(1)}(\mu)(X-\mu) + \frac{H^{(2)}(\mu)}{2}(X-\mu)^2 + R_1.$$

Aplicando a esperança, nós temos que

$$\begin{aligned}\mathbb{E}(Y) = \mathbb{E}\left\{H(\mu)\right\} + \mathbb{E}\left\{H^{(1)}(\mu)(X-\mu)\right\} \\ + \mathbb{E}\left\{\frac{H^{(2)}(\mu)}{2}(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\}.\end{aligned}$$

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$Y = H(\mu) + H^{(1)}(\mu)(X-\mu) + \frac{H^{(2)}(\mu)}{2}(X-\mu)^2 + R_1.$$

Aplicando a esperança, nós temos que

$$\begin{aligned}\mathbb{E}(Y) &= \mathbb{E}\left\{H(\mu)\right\} + \mathbb{E}\left\{H^{(1)}(\mu)(X-\mu)\right\} \\ &\quad + \mathbb{E}\left\{\frac{H^{(2)}(\mu)}{2}(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\}.\end{aligned}$$

Expressões aproximadas da média e da variância

Agora,

$$
\begin{aligned}
\mathbb{E}(Y) &= H(\mu) + H^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\} + \frac{H^{(2)}(\mu)}{2}\mathbb{E}\left\{(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\} \\
&= H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 + \mathbb{E}\left\{R_1\right\}.
\end{aligned}
$$

Expressões aproximadas da média e da variância

Agora,

$$\begin{aligned}\mathbb{E}(Y) &= H(\mu) + H^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\} + \frac{H^{(2)}(\mu)}{2}\mathbb{E}\left\{(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\}\\ &= H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 + \mathbb{E}\left\{R_1\right\}.\end{aligned}$$

Ignorando o termo $\mathbb{E}\left\{R_1\right\}$,

Expressões aproximadas da média e da variância

Agora,

$$\begin{aligned}\mathbb{E}(Y) &= H(\mu) + H^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\} + \frac{H^{(2)}(\mu)}{2}\mathbb{E}\left\{(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\} \\ &= H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 + \mathbb{E}\left\{R_1\right\}.\end{aligned}$$

Ignorando o termo $\mathbb{E}\left\{R_1\right\}$, nós temos que,

Expressões aproximadas da média e da variância

Agora,

$$\begin{aligned}\mathbb{E}(Y) &= H(\mu) + H^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\} + \frac{H^{(2)}(\mu)}{2}\mathbb{E}\left\{(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\} \\ &= H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 + \mathbb{E}\left\{R_1\right\}.\end{aligned}$$

Ignorando o termo $\mathbb{E}\left\{R_1\right\}$, nós temos que,

$$\mathbb{E}(Y) \approx H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2.$$

Expressões aproximadas da média e da variância

Agora,

$$\begin{aligned}\mathbb{E}(Y) &= H(\mu) + H^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\} + \frac{H^{(2)}(\mu)}{2}\mathbb{E}\left\{(X-\mu)^2\right\} + \mathbb{E}\left\{R_1\right\} \\ &= H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2 + \mathbb{E}\left\{R_1\right\}.\end{aligned}$$

Ignorando o termo $\mathbb{E}\left\{R_1\right\}$, nós temos que,

$$\mathbb{E}(Y) \approx H(\mu) + \frac{H^{(2)}(\mu)}{2}\sigma^2.$$

Expressões aproximadas da média e da variância

Demonstração 2. Usando (1),

Expressões aproximadas da média e da variância

Demonstração 2. Usando (1), até o primeiro termo

Expressões aproximadas da média e da variância

Demonstração 2. Usando (1), até o primeiro termo e em torno de $X = \mu$,

Expressões aproximadas da média e da variância

Demonstração 2. Usando (1), até o primeiro termo e em torno de $X = \mu$,

nós temos que

Demonstração 2. Usando (1), até o primeiro termo e em torno de $X = \mu$, nós temos que

$$
\begin{aligned}
Y = H(X) &= \sum_{n=0}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n \\
&= \frac{H^{(0)}(\mu)}{0!}(X-\mu)^0 + \frac{H^{(1)}(\mu)}{1!}(X-\mu)^1 \\
&\quad + \underbrace{\sum_{n=2}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n}_{R_2}.
\end{aligned}
$$

Expressões aproximadas da média e da variância

Demonstração 2. Usando (1), até o primeiro termo e em torno de $X = \mu$, nós temos que

$$
\begin{aligned}
Y = H(X) &= \sum_{n=0}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n \\
&= \frac{H^{(0)}(\mu)}{0!}(X-\mu)^0 + \frac{H^{(1)}(\mu)}{1!}(X-\mu)^1 \\
&\quad + \underbrace{\sum_{n=2}^{\infty} \frac{H^{(n)}(\mu)}{n!}(X-\mu)^n}_{R_2}.
\end{aligned}
$$

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$
\begin{aligned}
& Y = H(\mu) + H^{(1)}(\mu)(X - \mu) + R_2 \\
\Rightarrow\ & Y \approx H(\mu) + H^{(1)}(\mu)(X - \mu).
\end{aligned}
$$

Simplificando as expressões anteriores, nós temos que

$$Y = H(\mu) + H^{(1)}(\mu)(X - \mu) + R_2$$
$$\Rightarrow Y \approx H(\mu) + H^{(1)}(\mu)(X - \mu).$$

Aplicando a variância,

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$Y = H(\mu) + H^{(1)}(\mu)(X-\mu) + R_2$$
$$\Rightarrow Y \approx H(\mu) + H^{(1)}(\mu)(X-\mu).$$

Aplicando a variância, nós temos que

$$\begin{aligned}\mathrm{Var}(Y) &\approx \mathrm{Var}\left\{H(\mu) + H^{(1)}(\mu)(X-\mu)\right\}\\ &\approx \left[H^{(1)}(\mu)\right]^2 \mathrm{Var}(X-\mu) = \left[H^{(1)}(\mu)\right]^2 \sigma^2.\end{aligned}$$

Expressões aproximadas da média e da variância

Simplificando as expressões anteriores, nós temos que

$$
\begin{aligned}
& Y = H(\mu) + H^{(1)}(\mu)(X - \mu) + R_2 \\
\Rightarrow\ & Y \approx H(\mu) + H^{(1)}(\mu)(X - \mu).
\end{aligned}
$$

Aplicando a variância, nós temos que

$$
\begin{aligned}
\mathrm{Var}(Y) &\approx \mathrm{Var}\left\{H(\mu) + H^{(1)}(\mu)(X - \mu)\right\} \\
&\approx \left[H^{(1)}(\mu)\right]^2 \mathrm{Var}(X - \mu) = \left[H^{(1)}(\mu)\right]^2 \sigma^2.
\end{aligned}
$$

Roteiro

Desigualdade de Markov

Se X é uma variável aleatória que assume somente valores não negativos,

então para qualquer $\lambda > 0$,

$$\mathbb{P}\left(X \geq \lambda\right) \leq \frac{\mathbb{E}\left(X\right)}{\lambda}.$$

Desigualdade de Markov

Se X é uma variável aleatória que assume somente valores não negativos, então para qualquer $\lambda > 0$,

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}.$$

Desigualdade de Markov

Prova. Para $\lambda > 0$,

Desigualdade de Markov

Prova. Para $\lambda > 0$, seja

Desigualdade de Markov

Prova. Para $\lambda > 0$, seja

$$I = \begin{cases} 1, & \text{se } X \geq \lambda; \\ 0, & \text{caso contrário.} \end{cases}$$

Desigualdade de Markov

Prova. Para $\lambda > 0$, seja

$$I = \begin{cases} 1, & \text{se } X \geq \lambda; \\ 0, & \text{caso contrário.} \end{cases}$$

Notem que,

Desigualdade de Markov

Prova. Para $\lambda > 0$, seja

$$I = \begin{cases} 1, & \text{se } X \geq \lambda; \\ 0, & \text{caso contrário.} \end{cases}$$

Notem que, $\mathbb{E}(I) = \mathbb{P}(X \geq \lambda)$.

Desigualdade de Markov

Prova. Para $\lambda > 0$, seja

$$I = \begin{cases} 1, & \text{se } X \geq \lambda; \\ 0, & \text{caso contrário.} \end{cases}$$

Notem que, $\mathbb{E}(I) = \mathbb{P}(X \geq \lambda)$.

Desigualdade de Markov

Notem também que,

$$I \leq \frac{X}{\lambda}.$$

Desigualdade de Markov

Notem também que,

$$I \leq \frac{X}{\lambda}.$$

Tomando as esperanças da desigualdade acima,

Desigualdade de Markov

Notem também que,

$$I \leq \frac{X}{\lambda}.$$

Tomando as esperanças da desigualdade acima,

$$\mathbb{E}(I) \leq \mathbb{E}\left(\frac{X}{\lambda}\right) = \frac{\mathbb{E}(X)}{\lambda}$$

$$\Rightarrow \mathbb{P}\left(X \geq \lambda\right) \leq \frac{\mathbb{E}\left(X\right)}{\lambda}.$$

Desigualdade de Markov

Notem também que,

$$I \leq \frac{X}{\lambda}.$$

Tomando as esperanças da desigualdade acima,

$$\mathbb{E}(I) \leq \mathbb{E}\left(\frac{X}{\lambda}\right) = \frac{\mathbb{E}(X)}{\lambda}$$

$$\Rightarrow \mathbb{P}\left(X \geq \lambda\right) \leq \frac{\mathbb{E}\left(X\right)}{\lambda}.$$

Desigualdade de Chebyshev

Seja X uma variável aleatória com média $\mu < \infty$ (finita) e variância σ^2.

Então, para todo $\lambda > 0$,

$$\mathbb{P}\left(|X - \mu| \geq \lambda\right) \leq \frac{\sigma^2}{\lambda^2}.$$

Desigualdade de Chebyshev

Seja X uma variável aleatória com média $\mu < \infty$ (finita) e variância σ^2.

Então, para todo $\lambda > 0$,

$$\mathbb{P}\left(|X - \mu| \geq \lambda\right) \leq \frac{\sigma^2}{\lambda^2}.$$

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa,

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

Prova. Desde que $(X-\mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X-\mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2}.$$

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X-\mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2}.$$

Observem que,

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X - \mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X - \mu)^2]}{\lambda^2}.$$

Observem que, $(X - \mu)^2 \geq \lambda^2$

Desigualdade de Chebyshev

Prova. Desde que $(X-\mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X-\mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2}.$$

Observem que, $(X-\mu)^2 \geq \lambda^2$ se e somente se $|X-\mu| \geq \lambda$,

Prova. Desde que $(X-\mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X-\mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2}.$$

Observem que, $(X-\mu)^2 \geq \lambda^2$ se e somente se $|X-\mu| \geq \lambda$, dessa forma,

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X - \mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X - \mu)^2]}{\lambda^2}.$$

Observem que, $(X - \mu)^2 \geq \lambda^2$ se e somente se $|X - \mu| \geq \lambda$, dessa forma, a equação acima é equivalente a

Desigualdade de Chebyshev

Prova. Desde que $(X-\mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X-\mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2}.$$

Observem que, $(X-\mu)^2 \geq \lambda^2$ se e somente se $|X-\mu| \geq \lambda$, dessa forma, a equação acima é equivalente a

$$\mathbb{P}\left(|X-\mu| \geq \lambda\right) \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2} = \frac{\sigma^2}{\lambda^2}.$$

Desigualdade de Chebyshev

Prova. Desde que $(X - \mu)^2$ é uma variável aleatória não negativa, nós podemos aplicar a desigualdade de Markov para obter

$$\mathbb{P}\left[(X-\mu)^2 \geq \lambda^2\right] \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2}.$$

Observem que, $(X-\mu)^2 \geq \lambda^2$ se e somente se $|X-\mu| \geq \lambda$, dessa forma, a equação acima é equivalente a

$$\mathbb{P}\left(|X-\mu| \geq \lambda\right) \leq \frac{\mathbb{E}[(X-\mu)^2]}{\lambda^2} = \frac{\sigma^2}{\lambda^2}.$$

Desigualdades

A importância das desigualdades de Markov e Chebyshev é que elas nos permitem encontrar limites para as probabilidades quando **somente** a média ou ambas, a média e variância, da distribuição de probabilidades são conhecidas.

Desigualdades

Claro, se a distribuição é conhecida, então a probabilidade desejada pode ser calculada de forma exata e nós não precisaríamos recorrer a esses limites (essas desigualdades).

Desigualdade de Markov generalizada

Seja X uma variável aleatória qualquer. Para todo $t > 0$,

$$\mathbb{P}\left(|X| \geq \lambda\right) \leq \frac{\mathbb{E}\left[|X|^t\right]}{\lambda^t},\ \forall \lambda > 0.$$

Desigualdade de Markov generalizada

Seja X uma variável aleatória qualquer. Para todo $t > 0$,

$$\mathbb{P}\left(|X| \geq \lambda\right) \leq \frac{\mathbb{E}\left[|X|^t\right]}{\lambda^t}, \ \forall \lambda > 0.$$

Limitantes de Chernoff

Para quaisquer variável aleatória X e constante $a \in \mathbb{R}$,

$$\mathbb{P}\left(X \geq a\right) \leq e^{-ta}\, M_X(t),\ \forall t > 0;$$

$$\mathbb{P}\left(X \leq a\right) \leq e^{-ta}\, M_X(t),\ \forall t < 0;$$

Limitantes de Chernoff

Para quaisquer variável aleatória X e constante $a \in \mathbb{R}$,

$$\mathbb{P}\,(X \geq a) \leq e^{-ta}\, M_X(t),\ \forall t > 0;$$

$$\mathbb{P}\,(X \leq a) \leq e^{-ta}\, M_X(t),\ \forall t < 0;$$

em que $M_X(t)$ é a função geradora de momentos de X, avaliada no ponto t.

Limitantes de Chernoff

Para quaisquer variável aleatória X e constante $a \in \mathbb{R}$,

$$\mathbb{P}\,(X \geq a) \leq e^{-ta}\, M_X(t),\ \forall t > 0;$$

$$\mathbb{P}\,(X \leq a) \leq e^{-ta}\, M_X(t),\ \forall t < 0;$$

em que $M_X(t)$ é a função geradora de momentos de X, avaliada no ponto t. Uma vez que os limitantes de Chernoff valem para $t \in \mathbb{R}$,

Limitantes de Chernoff

Para quaisquer variável aleatória X e constante $a \in \mathbb{R}$,

$$\mathbb{P}\left(X \geq a\right) \leq e^{-ta} M_X(t), \ \forall t > 0;$$

$$\mathbb{P}\left(X \leq a\right) \leq e^{-ta} M_X(t), \ \forall t < 0;$$

em que $M_X(t)$ é a função geradora de momentos de X, avaliada no ponto t. Uma vez que os limitantes de Chernoff valem para $t \in \mathbb{R}$, nós obtemos o melhor limitante para $\mathbb{P}\left(X \geq a\right)$ escolhendo t que minimiza $e^{-ta} M_X(t)$.

Limitantes de Chernoff

Para quaisquer variável aleatória X e constante $a \in \mathbb{R}$,

$$\mathbb{P}\left(X \geq a\right) \leq e^{-ta} M_X(t), \ \forall t > 0;$$

$$\mathbb{P}\left(X \leq a\right) \leq e^{-ta} M_X(t), \ \forall t < 0;$$

em que $M_X(t)$ é a função geradora de momentos de X, avaliada no ponto t. Uma vez que os limitantes de Chernoff valem para $t \in \mathbb{R}$, nós obtemos o melhor limitante para $\mathbb{P}\left(X \geq a\right)$ escolhendo t que minimiza $e^{-ta} M_X(t)$.

Função convexa

Definição

Uma função real e duas vezes diferenciável $f(a)$ é dita ser convexa se $f^{(2)}(a) \geq 0$, para todo a.

Função convexa

Definição

Uma função real e duas vezes diferenciável $f(a)$ é dita ser convexa se $f^{(2)}(a) \geq 0$, para todo a. Além disso, suas derivadas são monotonicamente não decrescentes.

Função convexa

Definição

Uma função real e duas vezes diferenciável $f(a)$ é dita ser convexa se $f^{(2)}(a) \geq 0$, para todo a. Além disso, suas derivadas são monotonicamente não decrescentes.

Observação. De forma similar, $f(a)$ é dita ser côncava se $f^{(2)}(a) \leq 0$.

Função convexa

Definição

Uma função real e duas vezes diferenciável $f(a)$ é dita ser convexa se $f^{(2)}(a) \geq 0$, para todo a. Além disso, suas derivadas são monotonicamente não decrescentes.

Observação. De forma similar, $f(a)$ é dita ser côncava se $f^{(2)}(a) \leq 0$.

Desigualdade de Jensen

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função convexa. Se a variável aleatória X é integrável,

Desigualdade de Jensen

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função convexa. Se a variável aleatória X é integrável,

então

$$\mathbb{E}\left[f\left(X\right)\right] \geq f\left(\mathbb{E}\left[X\right]\right).$$

Desigualdade de Jensen

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função convexa. Se a variável aleatória X é integrável, então

$$\mathbb{E}\left[f\left(X\right)\right] \geq f\left(\mathbb{E}\left[X\right]\right).$$

Observação: A Desigualdade de Jensen é válida se f é convexa em um intervalo (a, b) tal que $\mathbb{P}\left(a < X < b\right) = 1$, em que se admite a possibilidade de $a = -\infty$ ou $b = +\infty$.

Desigualdade de Jensen

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função convexa. Se a variável aleatória X é integrável, então

$$\mathbb{E}\left[f\left(X\right)\right] \geq f\left(\mathbb{E}\left[X\right]\right).$$

Observação: A Desigualdade de Jensen é válida se f é convexa em um intervalo (a, b) tal que $\mathbb{P}\left(a < X < b\right) = 1$, em que se admite a possibilidade de $a = -\infty$ ou $b = +\infty$.

Desigualdade de Jensen

Demonstração. Usando (2),

Desigualdade de Jensen

Demonstração. Usando (2), até o segundo termo

Desigualdade de Jensen

Demonstração. Usando (2), até o segundo termo e em torno de $X = \mu$,

Desigualdade de Jensen

Demonstração. Usando (2), até o segundo termo e em torno de $X = \mu$,

nós temos que

Desigualdade de Jensen

Demonstração. Usando (2), até o segundo termo e em torno de $X = \mu$, nós temos que

$$f(X) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\mu)}{n!}(X-\mu)^n$$

$$= \frac{f^{(0)}(\mu)}{0!}(X-\mu)^0 + \frac{f^{(1)}(\mu)}{1!}(X-\mu)^1 + \underbrace{\sum_{n=2}^{\infty} \frac{f^{(n)}(\mu)}{n!}(X-\mu)^n}_{R_2(x)}.$$

Demonstração. Usando (2), até o segundo termo e em torno de $X = \mu$, nós temos que

$$f(X) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\mu)}{n!}(X-\mu)^n$$

$$= \frac{f^{(0)}(\mu)}{0!}(X-\mu)^0 + \frac{f^{(1)}(\mu)}{1!}(X-\mu)^1 + \underbrace{\sum_{n=2}^{\infty} \frac{f^{(n)}(\mu)}{n!}(X-\mu)^n}_{R_2(x)}.$$

Desigualdade de Jensen

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$,

Desigualdade de Jensen

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$,

nós temos que

Desigualdade de Jensen

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$, nós temos que

$$f(X) = f(\mu) + f^{(1)}(\mu)(X-\mu) + \frac{f^{(2)}(\tau)}{2}(X-\mu)^2,$$

Desigualdade de Jensen

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$,

nós temos que

$$f(X) = f(\mu) + f^{(1)}(\mu)(X - \mu) + \frac{f^{(2)}(\tau)}{2}(X - \mu)^2,$$

em que $X \leq \tau \leq \mu$.

Desigualdade de Jensen

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$, nós temos que

$$f(X) = f(\mu) + f^{(1)}(\mu)(X - \mu) + \frac{f^{(2)}(\tau)}{2}(X - \mu)^2,$$

em que $X \leq \tau \leq \mu$. Como f é uma função convexa,

Desigualdade de Jensen

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$, nós temos que

$$f(X) = f(\mu) + f^{(1)}(\mu)(X-\mu) + \frac{f^{(2)}(\tau)}{2}(X-\mu)^2,$$

em que $X \leq \tau \leq \mu$. Como f é uma função convexa, nós temos que

$$f(X) \geq f(\mu) + f^{(1)}(\mu)(X-\mu).$$

Simplificando as expressões anteriores e utilizando o resultado para $R_2(x)$, nós temos que

$$f(X) = f(\mu) + f^{(1)}(\mu)(X - \mu) + \frac{f^{(2)}(\tau)}{2}(X - \mu)^2,$$

em que $X \leq \tau \leq \mu$. Como f é uma função convexa, nós temos que

$$f(X) \geq f(\mu) + f^{(1)}(\mu)(X - \mu).$$

Desigualdade de Jensen

Aplicando a esperança, nós temos que

$$\mathbb{E}\left[f\left(X\right)\right] \geq \mathbb{E}\left\{f(\mu)\right\} + \mathbb{E}\left\{f^{(1)}(\mu)(X-\mu)\right\}.$$

Desigualdade de Jensen

Aplicando a esperança, nós temos que

$$\mathbb{E}\left[f\left(X\right)\right] \geq \mathbb{E}\left\{f(\mu)\right\} + \mathbb{E}\left\{f^{(1)}(\mu)(X-\mu)\right\}.$$

Agora,

Desigualdade de Jensen

Aplicando a esperança, nós temos que

$$\mathbb{E}\left[f\left(X\right)\right] \geq \mathbb{E}\left\{f(\mu)\right\} + \mathbb{E}\left\{f^{(1)}(\mu)(X-\mu)\right\}.$$

Agora,

$$\mathbb{E}\left[f\left(X\right)\right] \geq f(\mu) + f^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\}$$

$$\Rightarrow \mathbb{E}\left[f\left(X\right)\right] \geq f(\mu).$$

Desigualdade de Jensen

Aplicando a esperança, nós temos que

$$\mathbb{E}\left[f\left(X\right)\right] \geq \mathbb{E}\left\{f(\mu)\right\} + \mathbb{E}\left\{f^{(1)}(\mu)(X-\mu)\right\}.$$

Agora,

$$\mathbb{E}\left[f\left(X\right)\right] \geq f(\mu) + f^{(1)}(\mu)\mathbb{E}\left\{X-\mu\right\}$$

$$\Rightarrow \mathbb{E}\left[f\left(X\right)\right] \geq f(\mu).$$

Desigualdade de Jensen

Lembrando que $\mathbb{E}(X) = \mu$, nós concluímos a demonstração.

Desigualdade de Jensen

Lembrando que $\mathbb{E}(X) = \mu$, nós concluímos a demonstração.

Essa demonstração foi adaptada de Ross (2019, p. 427),

Desigualdade de Jensen

Lembrando que $\mathbb{E}(X) = \mu$, nós concluímos a demonstração.

Essa demonstração foi adaptada de Ross (2019, p. 427), para uma forma alternativa,

Desigualdade de Jensen

Lembrando que $\mathbb{E}(X) = \mu$, nós concluímos a demonstração.

Essa demonstração foi adaptada de Ross (2019, p. 427), para uma forma alternativa, ver Magalhães (2015, p. 224).

Desigualdade de Jensen

Lembrando que $\mathbb{E}(X) = \mu$, nós concluímos a demonstração.

Essa demonstração foi adaptada de Ross (2019, p. 427), para uma forma alternativa, ver Magalhães (2015, p. 224).

Desigualdade de Cauchy-Schwarz

Sejam X e Y variáveis aleatórias com esperanças finitas. Então

$$\mathbb{E}\left(|XY|\right) \leq \sqrt{\mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)}.$$

Desigualdade de Cauchy-Schwarz

Sejam X e Y variáveis aleatórias com esperanças finitas. Então

$$\mathbb{E}\left(|XY|\right) \leq \sqrt{\mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)}.$$

Roteiro

Referências bibliográficas I

Magalhães, M. N. (2015), *Probabilidade e variáveis aleatórias*, 3 edn, Editora da Universidade de São Paulo, São Paulo.

Ross, S. (2019), *A first course in probability*, 10 edn. global edn, Pearson Education, Malaysia.

Taylor, B. (1715), *Methodus Incrementorum Directa et Inversa*, Typis Pearsonianis prostant apud Gul. Innys ad Insignia Principis in Coemeterio Paulino, Londini.

Obrigado!

✉ tiago.magalhaes@ufjf.br

ufjf.br/tiago_magalhaes

Departamento de Estatística, Sala 319