Lista 3 Probabilidade 2020 Maio 20

Nestor Caticha

May 2020

Exercício 1. A variável X_1 tem distribuição densidade de probabilidade dada pela função $f(x)$. A variável X_2 tem densidade condicional $P(X_2|X1 = x)$ uniforme no intervalo $[0, x]$

1. Encontre uma expressão para $P(X_2|I)$ dada a informação do enunciado
2. Encontre $P(X_2|I)$ no caso que $f(x) = \frac{1}{x^2}$

Dica 1: Escreva as densidades acima usando a função de Heaviside $\Theta(x - a)$ que é 1 se $x > a$ e zero se não. Dica 2 use as regras do produto e soma.

Exercício 2. No problema de urnas com reposição precisamos calcular a integral de Euler

$$E_k^r = \int_0^1 p^r(1-p)^k dp \quad = \quad \frac{r!k!}{(r+k+1)!} \tag{1}$$

Mostre este resultado.

Exercício 3. X_1 é uma variável aleatória com densidade de probabilidade $f(x)$ no intervalo $[0, 1]$. Um valor de X_1 é escolhido, chame seu valor de x_p. Uma moeda é jogada n vezes com probabilidade x_p de sair cara em cada jogada. Calcule o valor esperado do número de caras k nas n jogadas

1. As jogadas são independentes e o valor de x_p é o mesmo em cada jogada. Encontre uma expressão para $\mathbb{E}(k)$ no caso de $f(x)$ geral
2. Encontre $\mathbb{E}(k)$ no caso de $f(x)$ ser uniforme em $[0, 1]$
3. Encontre $\mathbb{E}(k)$ no caso de $f(x)$ ser uniforme em $[0, 1]$ e um novo valor de x_p é sorteado para cada jogada da moeda.

Você vai precisar em algum lugar da integral de Euler do exercício anterior.

Exercício 4. Duas variáveis aleatórias independentes tem densidades uniformes: X_1 em $[0, A]$ e X_2 em $[0, B]$ com $A < B$. Encontre a probabilidade de "$X_1 < X_2$ ". Dica faça um desenho no plano $x - 1, x_2$.

Exercício 5. Suponha N variáveis aleatórias independentes e igualmente distribuidas com distribuição exponencial $P(X_i|a) = \frac{1}{a}e^{-\frac{x}{a}}$. Chame a soma de $Y_N = \sum_{i=1}^{N} X_i$

1. Encontre a distribuição de Y_N exatamente.
2. Compare os valores médios e variância de Y_N com o esperado pela teorema do limite central.

Os exercícios de 6-9 estão no meio do texto que segue. Este é complementar às notas. Traz resultados novos para vocês. Os exercícios são simples e estão distribuidos dentro dotexto. Apresentamos novamente a desigualdade de Chebyshev, obtida de uma forma mais geral que o que fizemos em aula. O objetivo, além de apresentar as desigualdades, é obter a Lei Fraca dos Grandes Números.

0.1 Introdução: Algumas desigualdades

Raramente ao lidar com modelos em Física que levam a problemas em probabilidades podemos fazer as contas de forma exata. Para progredir há algumas estratégias possíveis. Uma forma de aproximação é fazer a aproximação no modelo de interesse e depois tratar exatamente o modelo simplificado. Talvez seja estranho dizer que isto é um método de aproximar, pois as contas são exatas. Sim, exatas mas no modelo menos realista. Outra, que é encontrada em muitos dos métodos da Física Teórica, é fazer aproximações nas expressões associadas ao modelo original, que permitam avançar analiticamente e chegar a alguma conclusão. As aproximações tem a vantagem que muitas vezes, se bem sucedidas, capturam a essência do que um modelo tem que leva a algum fenômeno de interesse. A desvantagem é que ao fazer aproximações não sabemos se uma quantidade de interesse está sendo super ou subestimada. Outra estratégia é usar computadores e isso é a base da Física Computacional. Veremos um pouco disto em outros capítulos, em particular ao lidar com o método de Monte Carlo. Uma terceira estrategia usa a ideia de desigualdades, que obviamente tem sua origem em Análise. De forma simplificada são a essência de métodos rigorosos de Física Matemática.
Há uma literatura enorme sobre desigualdades e não temos nenhuma possibilidade de tratar o problema de forma exaustiva. Simplesmente queremos mostrar aos alunos alguns resultados simples e rigorosos.

0.2 Desigualdade de Jensen

Damos a seguir uma versão muito limitada de uma desigualdade que serve em muitos contextos. A função e^x tem curvatura sempre com o mesmo sinal ($d^2e^x/d^2x > 0$ para qualquer x). Funções com esta característica são ditas convexas. Verifique que a função $-\log(x)$ também é convexa para qualquer $x > 0$. O conceito pode ser estendido a mais dimensões. Aqui nos restringimos à função exponencial. Suponha uma variável aleatória x com densidade $p(x)$. Então vale

$$\mathbb{E}(e^x) \geq e^{\mathbb{E}(x)} \tag{2}$$

Começamos com

$$\mathbb{E}(e^x) = \mathbb{E}(e^{x+\mathbb{E}(x)-\mathbb{E}(x)}) = e^{\mathbb{E}(x)}\mathbb{E}(e^{x-\mathbb{E}(x)}) \tag{3}$$

e no segundo fator usamos a desigualdade óbvia (verifique)

$$e^z \geq 1 + z, \tag{4}$$

com $z = x - \mathbb{E}(x)$, portanto

$$\mathbb{E}(e^x) = e^{\mathbb{E}(x)}\mathbb{E}(e^{x-\mathbb{E}(x)}) \geq e^{\mathbb{E}(x)}\mathbb{E}(1 + x - \mathbb{E}(x)) = e^{\mathbb{E}(x)}(1 + \mathbb{E}(x) - \mathbb{E}(x)), \tag{5}$$

que leva à desigualdade 2. O valor da média da função exponencial (convexa) é maior que o valor da função exponencial calculada no valor médio da variável aleatória. Isto é devido ao simples fato que a curvatura é positiva. Portanto não é devido a nada além da convexidade da exponencial e por isso vale para outras funções convexas $g(x)$

$$\mathbb{E}(g(x)) \geq g(\mathbb{E}(x)). \tag{6}$$

Aqui só usaremos o caso da exponencial que foi provado e não o caso mais geral que não o foi.

Exercício 6.

1. Mostre a desigualdade 4

2. Refaça a dedução da desigualdade de Jensen
3. Esboce um gráfico (a mão) para mostrar intuitivamente a desigualdade de Jensen. Suponha que x tem uma distribuição uniforme no intervalo $x \in [a, b]$, $a < b$. Desenhe o gráfico de uma função convexa $f(x)$ nesse intervalo e marque no eixo x o valor de $\mathbb{E}(x)$. Mostre que o ponto no gráfico de coordenadas $(\mathbb{E}(x), \mathbb{E}(f(x))$ está acima do ponto $(\mathbb{E}(x), f(\mathbb{E}(x))$
4. Mostre a desigualdade de Jensen, na expressão 6

0.3 Markov

A próxima desigualdade é devida a Markov (talvez já fosse conhecida por Chebyshev, seu orientador). Seja X uma variável aleatória não negativa que toma valores $[0, B)$ onde B pode ser infinito. A informação sobre X nos leva à densidade $p(x)$. Para qualquer $y \geq 0$ temos

$$
\begin{aligned}
P(X \geq y) &= \int_y^B p(x)dx && (7) \\
&= \int_0^B \mathbb{I}(x \geq y)p(x)dx && (8)
\end{aligned}
$$

onde $\mathbb{I}(x \geq y)$ é a função indicadora do evento $x \geq y$. Isto é: $\mathbb{I}(x \geq y) = 1$ se $x \geq y$ e $\mathbb{I}(x \geq y) = 0$ se $x < y$. Note que se $x \geq y$, então $x/y \geq 1 = \mathbb{I}(x \geq y)$, e se $x < y$ também vale $x/y \geq 0 = I(x \geq y)$. Portanto sempre vale que $x/y \geq \mathbb{I}(x \geq y)$. Então podemos substituir a igualdade 8, que talvez não saibamos calcular (se soubessemos para que fazer isto?), por uma desigualdade

$$
\begin{aligned}
P(X \geq y) &= \int_0^B \mathbb{I}(x \geq y)p(x)dx && (9) \\
&\leq \int_0^B \frac{x}{y}p(x)dx && (10) \\
&= \frac{1}{y}\int_0^B xp(x)dx && (11) \\
P(X \geq y) &\leq \frac{\mathbb{E}(X)}{y}. && (12)
\end{aligned}
$$

Esta é a desigualdade de Markov. Substituimos uma igualdade, a eq. 8, talvez intratável analiticamente por uma desigualdade, que talvez nos dê algo concreto e rigoroso, embora possa não ser preciso. Talvez o estudante não esteja impressionado nesta altura. Mas, fazendo variações sobre o tema, este é o ponto de partida para várias outras desigualdades que muitas vezes representam a única forma de atacar rigorosamente alguns problemas.

Exercício 7. Aplique ao caso da distribuição exponencial. Mostre que $e^{-y} \leq \frac{1}{y}$, para $y \geq 0$.

0.4 Mudança de variáveis

Seja X como acima. Suponha que Z é uma função de X, i.e. toma valores z dados por $z = f(x)$ e tem densidade de probabilidade $\tilde{p}(z)$. Lembre que a relação entre as densidades é

$$p(x)dx = \tilde{p}(z)dz.$$

Partimos da desigualdade de Markov para a variável aleatória Z e mudamos variáveis

$$P(Z \geq y) \leq \frac{1}{y} \int_{f(0)}^{f(B)} z\tilde{p}(z)dz \tag{13}$$

$$= \frac{1}{y} \int_{0}^{B} f(x)p(x)dx \tag{14}$$

Portanto temos

$$P(f(x) \geq y) \leq \frac{I\!\!E(f(x))}{y} \tag{15}$$

onde $I\!\!E(f(x)) = \int f(x)p(x)dx$.

0.5 Chebyshev

Consideremos o caso particular $f(x) = (x - I\!\!E(X))^2$. Isto é interessante porque se X se afasta do valor esperado por mais que uma quantidade $t > 0$ temos duas possibilidades, ou

$$x \geq I\!\!E(x) + t$$

ou

$$x \leq I\!\!E(x) - t.$$

As duas condições podem ser escritas como $f(x) \geq t^2$. Portanto a desigualdade de Markov nos leva a

$$P((x - I\!\!E(X))^2 \geq t^2) \leq \frac{I\!\!E((x - I\!\!E(X))^2)}{t^2}$$

$$P((x - I\!\!E(X))^2 \geq t^2) \leq \frac{\mathrm{Var}(X)}{t^2} \tag{16}$$

Esta última é conhecida como desigualdade de Chebyshev. A probabilidade de que X tenha um valor que se afasta da média por mais de um valor $t > 0$ cai com t^2 e a escala é dada pela variância de X. É claro que se a variância for infinita a desigualdade não diz nada.

0.6 Lei Fraca dos grandes números

Há vários teoremas associados à lei dos grandes números. Olharemos os mais simples. De forma suscinta, considere uma sequência de variáveis aleatórias $X_1, X_2, ...X_n$ *independentes* com média finita e *variância finita*. Suponha que para qualquer n, $\exists M < \infty$,tal que as variâncias $\sigma_i^2 < M$ (ou seja $M < \infty$ é independente de n e as variâncias são uniformemente limitadas). Defina as somas parciais

$$W_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \tag{17}$$

que também formam uma sequência de variáveis aleatórias. Então W_n converge para um número dado pelo limite da sua média $\lim_{n\to\infty} I\!\!E(W_n)$. Mas o que significa se aproxima ou converge? Dependendo das condições sobre os X_i (e da quantidade de esforço a ser gasto na prova) diferentes medidas de distancia podem ser dadas. Para obter a lei Fraca dos Grandes números definimos a **convergência em probabilidade**:

$$\lim_{n\to\infty} Prob(|W_n - E(W_n)| \geq \epsilon) = 0. \tag{18}$$

Ou seja, W_n é uma variável aleatória, mas a probabilidade que se afaste mais que ϵ do seu valor esperado, vai a zero quando n aumenta. Dados um $\epsilon > 0$ e $\delta > 0$ existe um $n_0(\epsilon, \delta)$ tal que se $n > n_0$, então a probabilidade de $|W_n - \mathbb{E}(W_n)| < \delta$.

Exercício 8. Suponha que os X_i sejam independentes, igualmente distribuidos (iid) e $\mathbb{E}(x_i) = \mu$.. Use a desigualdade de Chebyshev e demonstre a Lei Fraca dos grandes números.

A versão Forte é obtida impondo outro tipo de convergência, chamada **convergência quase certa** (em inglés é *almost surely.*) Assim dependendo de hipóteses sobre X_i, e se por exemplo $\mathbb{E}(X_i) = \mu$ para todo i caso seja verdade que

$$Prob(\lim_{n\to\infty} W_n = \mu) = 1 \tag{19}$$

0.7 Cotas exponenciais: Chernoff

Lembremos a função geradora dos momentos. Tomemos como definição

$$\Phi(\xi) = \mathbb{E}(e^{\xi x})$$

onde é melhor usar a versão real, embora seja às vezes conveniente introduzir a versão com exponenciais imaginárias.
Suponha que novamente olhemos para a probabilidade de $z \geq t$. Exponenciando, para $\xi > 0$, temos $e^{\xi x} \geq e^{\xi t}$. Portanto, a equação 15 dá:

$$P(x \geq t) = P(e^{\xi x} \geq e^{\xi t}) \leq \frac{\mathbb{E}(e^{\xi x})}{e^{\xi t}}. \tag{20}$$

mas o valor de ξ está à nossa disposição e

$$P(x \geq t) \leq \min_{\xi}\{\mathbb{E}(e^{\xi x})e^{-\xi t}\}. \tag{21}$$

Para avançar mais é preciso fazer hipóteses sobre as propriedades de $P(x)$. Faremos alguns exemplos simples para ver quão boa ou ruim pode ser esta desigualdade.
Caso exponencial: $x \sim \exp(-x)$.
Os valores médios $\mathbb{E}(x) = 1$ e $\mathbb{E}(x^2) = 2$ são fáceis de calcular. Perguntamos qual a probabilidade que x seja maior que o valor médio por uma quantidade t. Note que é fácil fazer as contas:

$$P(x \geq 1+t) = \int_{1+t}^{\infty} e^{-x} dx = e^{-(1+t)},$$

mas suponha que não sabemos. Para $0 \leq \xi < 1$ temos

$$\mathbb{E}(e^{\xi x}) = \int_0^{\infty} e^{\xi x} e^{-x} dx = \frac{1}{1-\xi}$$

e

$$P(x \geq 1+t) = P(e^{\xi x} \geq e^{\xi(1+t)}) \leq \min_{\xi} \mathbb{E}(e^{\xi x})e^{-\xi(1+t)} = \min_{\xi} \frac{e^{-\xi(1+t)}}{1-\xi}.$$

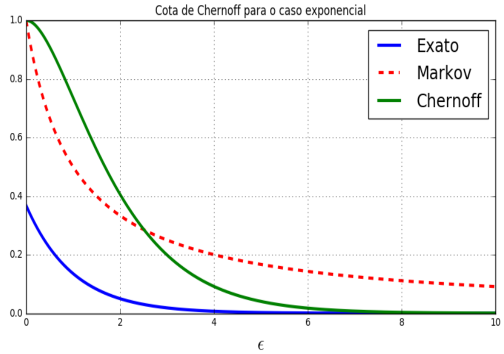

Figura 1: $P(x \geq \epsilon)$. Resultado exato e cotas de Markov de Chernoff para o caso exponencial

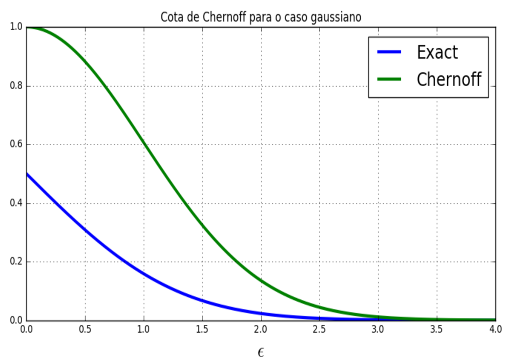

Figura 2: $P(x \geq \epsilon)$. Resultado exato e cota de Chernoff para o caso gaussiano

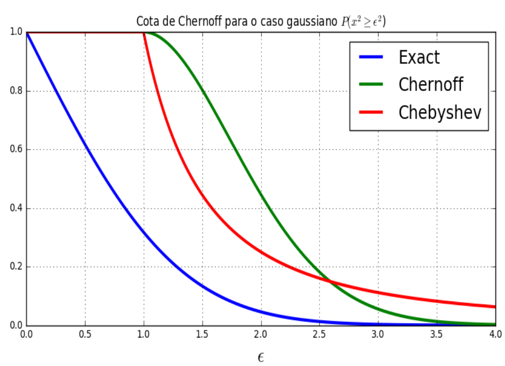

Figura 3: $P(x \geq \epsilon$ ou $x \leq -\epsilon)$. Resultado exato e cotas de Chebyshev e Chernoff para o caso gaussiano

O mínimo ocorre para $\xi = t/(1+t)$ e leva à cota de Chernoff

$$P(x \geq 1+t) \leq (1+t)e^{-t}$$

Caso Gaussiano: $x \sim \mathcal{N}(0,\sigma)$
A função geradora

$$\Phi(\xi) = I\!E(e^{\xi x}) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}+\xi x}dx$$

Completando quadrados

$$-\frac{x^2}{2\sigma^2} + \xi x + \frac{A^2 - A^2}{2\sigma^2} = -\frac{(x-A)^2}{2\sigma^2} + \frac{A^2}{2\sigma^2}$$

onde foi necessário escolher A para cancelar o termo linear ξx: $A = \xi\sigma^2$.

$$\Phi(\xi) = e^{\frac{A^2}{2\sigma^2}} = e^{\frac{\xi^2\sigma^2}{2}}$$

Note que na dedução da desigualdade de Markov usamos na desigualdade 10 que a variável aletória é não negativa para majorar a função indicadora. Aqui podemos fazer isto pois a exponencial é positiva. Voltando à equação 21

$$P(x \geq t) \leq \min_{\xi}\{e^{\frac{\xi^2\sigma^2}{2}}e^{-\xi t}\}. \tag{22}$$

O mínimo é obtido para $\xi = t/\sigma^2$ e portanto, colocando $t = \epsilon\sigma$

$$P(z \geq \epsilon\sigma) \leq e^{-\frac{\epsilon^2}{2}}. \tag{23}$$

Neste caso sabemos que

$$P(z \geq \epsilon\sigma) = \frac{1}{\sqrt{2\pi}}\int_{\epsilon}^{\infty} e^{-\frac{x^2}{2}}dx := \phi(\epsilon), \tag{24}$$

a função complementar da cumulativa.
Podemos perguntar qual é a cota para o caso $f(x) = (x - I\!E(X))^2$ que nos levou à desigualdade de Chebyshev. O valor exato

$$P(x \geq \epsilon \,\mathrm{ou}\, x \leq -\epsilon) = \frac{2}{\sqrt{2\pi}}\int_{\epsilon}^{\infty} e^{-\frac{x^2}{2}}dx := 2\phi(\epsilon). \tag{25}$$

Isto é igual a $P(e^{\xi^2 x^2} \geq e^{\xi^2\epsilon^2})$ que leva a uma desigualdade no estilo de Chernoff. Para $\xi < 1$:

$$\begin{aligned}
P(e^{\frac{1}{2}\xi^2x^2} \geq e^{\frac{1}{2}\xi^2\epsilon^2}) &= \int_{e^{\frac{1}{2}\xi^2x^2} \geq e^{\frac{1}{2}\xi^2\epsilon^2}}^{\infty} e^{-\frac{1}{2}x^2}\frac{dx}{\sqrt{2\pi}} && (26)\\
&\leq \min_{\xi}\int_{-\infty}^{\infty} e^{\frac{1}{2}\xi^2x^2}e^{-\frac{1}{2}\xi^2\epsilon^2}e^{-\frac{1}{2}x^2}\frac{dx}{\sqrt{2\pi}} && (27)\\
&= \min_{\xi}\frac{e^{-\frac{1}{2}\xi^2\epsilon^2}}{\sqrt{1-\xi^2}} && (28)
\end{aligned}$$

que é alcançado para $\xi^2 = 1 - 1/\epsilon^2$ quando $\epsilon \geq 1$ e $\xi = 1$ para $0 < \epsilon \leq 1$. Desta forma temos uma cota trivial para $\epsilon < 1$ e $P(x^2 \geq \epsilon^2) \leq \epsilon \exp((1-\epsilon^2)^2)$, mostrado na figura 3.

Os resultados obtidos acima não servem para outra coisa que desenvolver a intuição, pois não há sentido em usar cotas desta natureza quando as distribuições são tão simples que os cálculos podem ser feitos exatamente. O estilo de pensar sobre cotas requer um certo treinamento e será útil na prova de teoremas limite como o teorema central e em outros contextos.

Exercício 9. A variável aleatória tem distribuição exponencial: $P(X = x|a) = \frac{1}{a}e^{-\frac{x}{a}}$. Calcule a probabilidade $P(|X - \mu| > \epsilon\sigma)$ exatamente e compare com a cota de Chebyshev