RESUMO para a P2 de PROBABILIDADE - 2015

Por Arthur Salles. R$ 7,20

I - Média e Variância - variáveis aleatórias discretas

1) Para variáveis aleatórias discretas, define-se a média ou valor esperado como:

MÉDIA $$\boxed{\mu = E(x) = \sum_{j} x_j \cdot p(x_j)}$$

O que nada mais é do que uma média dos valores assumidos por x, ponderados por sua probabilidade de ocorrência.

Por exemplo:

x	2	5	10
p(x)	0,2	0,4	0,4

$E(x) = 2 \cdot 0{,}2 + 5 \cdot 0{,}4 + 10 \cdot 0{,}4$

$E(x) = 6{,}4$

2) Média de uma Função: se quisermos determinar o valor esperado não de x, mas de uma função de x h(x), podemos fazer:

VALOR ESPERADO DE UMA FUNÇÃO $$\boxed{E(h(x)) = \sum_{j} h(x_j) \cdot p(x_j)}$$

Por exemplo, se quiséssemos calcular a média de $h(x) = x^2 + 1$ no exercício anterior teríamos

x	2	5	10
$h(x)$	5	26	101
$p(x)$	0,2	0,4	0,4

$$E(h(x)) = \sum h(x) \cdot p(x)$$
$$E(h(x)) = 5 \cdot 0{,}2 + 26 \cdot 0{,}4 + 101 \cdot 0{,}4$$
$$E(h(x)) = 51{,}8$$

3) Propriedade da média : Se X é uma v.a discreta e a e b são números reais :

$$\boxed{E(ax+b) = a \cdot E(x) + b}$$

4) Variância : se X é uma v.a discreta com função de probabilidade $p(x)$, a variância de X é dada por:

VARIÂNCIA $$\boxed{var(X) = E((X-\mu)^2)} = \sum_j (x_j - \mu)^2 \cdot p(x_j)$$

Ou seja, é a média da distância ao quadrado dos valores que x assume até a sua média. Podemos interpretar isso como o quão "espalhados" estão os valores de X : se $var(x)$ é pequena, os valores se concentram perto da média ; se é grande, eles estão bem espalhados distantes da média

Uma outra forma de representar a variância é da forma σ^2, em que σ é o desvio padrão de X

DESVIO PADRÃO $\boxed{\sigma = \sqrt{var(X)}}$

Uma forma muito útil de calcular a variância é usar:

$$\boxed{var(X) = E(X^2) - [E(X)]^2}$$

5) Propriedade da variância :

$$\boxed{var(aX+b) = a^2 \cdot var(X)}$$

6) Variância de uma Função : se quisermos calcular a variância de uma função $h(x)$ fazemos

$$\boxed{var(h(x)) = \sum_j [h(x_j) - E(h(x))]^2 \cdot p(x_j)}$$

7) Propriedade para v.a. independentes : Se $X_1, X_2, \ldots X_m$ são variáveis aleatórias independentes, das quais conhecemos média $(E(x_j) = \mu_j)$ e variância $(Var(x_j) = \sigma_j^2)$ com $a_1, \ldots, a_m \in \mathbb{R}$:

i) $E(\sum a_j X_j) = \sum a_j E(X_j)$

ii) $Var(\sum a_j X_j) = \sum a_j^2 \, var(X_j)$

II - Média e Variância - variáveis aleatórias contínuas

Se X for uma variável aleatória contínua e se $f(x)$ for sua função densidade de probabilidade, podemos calcular:

1) Média: a média (ou esperança, ou valor esperado) de X é dado por:

MÉDIA $$\boxed{\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx}$$

2) Valor esperado de uma função: de forma similar ao que ocorre com as v.a. discretas, temos:

VALOR ESPERADO DE UMA FUNÇÃO $$\boxed{E(h(x)) = \int_{-\infty}^{\infty} h(x) \cdot f(x)\, dx}$$

3) Variância: a variância de X é dada por:

$$\boxed{\sigma^2 = Var(X) = E((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\, dx}$$

. Novamente, podemos calcular:

$$Var(X) = E(X^2) - [E(X)]^2$$

4) Propriedades:

Valem as mesmas propriedades que vimos para v.a. discretas:

i) $E(ax+b) = aE(x)+b$

ii) $Var(ax+b) = a^2 Var(x)$

Além disso, vale a propriedade do item 7, a respeito de variáveis aleatórias independentes.

III - Distribuições de Probabilidade Discretas

Existem algumas distribuições de probabilidade ($p(x)$) que aparecem com frequência em problemas reais. Por isso, elas recebem nomes bonitos e se tornam merecedoras de um estudo mais detalhado. Vejamos algumas delas:

1) Distribuição de Bernoulli: ocorre quando uma v.a. X assume apenas valores 1 ou 0, com probabilidade p e $(1-p)$, respectivamente. Temos:

$$\boxed{\begin{array}{l} p(x) \begin{cases} p(0) = 1-p \\ p(1) = p \end{cases} \\ \\ E(X) = p \\ var(X) = p(1-p) \end{array}}$$

2) Distribuição Binomial: suponha que façamos n ensaios de Bernoulli, em que cada ensaio é independente e com mesma probabilidade de sucesso p.

Se X é o número de sucessos nos n ensaios dizemos que X tem distribuição binomial e:

$$X = B(n\,;\,p)$$

$\uparrow$ ensaios $\nwarrow$ probabilidade sucesso

Esse tipo de distribuição é muito comum. Por exemplo se estou analisando 100 bandejas do bandeco e queremos descobrir quantas não têm resquícios do almoço anterior, estou fazendo um experimento binomial, em que um sucesso é pegar uma bandeja sem resquícios.

<u>Alguns valores importantes:</u>

$$P(X=k) = \binom{n}{k} p^{k} (1-p)^{n-k}$$

$$E(x) = np$$

$$var(x) = np(1-p)$$

Uma <u>propriedade</u> importante se relaciona com o resultado da soma de binomiais:

Se $X_1 = B(n_1, p)$, $X_2 = B(n_2, p)$, ..., $X_m = B(n_m, p)$

com $X_1, X_2, \dots, X_n$ variáveis <u>independentes</u>:

$$\sum_{i=1}^{m} X_i = B\left(\sum_{i=1}^{m} n_i, p\right)$$

3) Distribuição de Poisson:

Dada uma taxa de ocorrência de um fenômeno (λ), X é o número de ocorrências do fenômeno em um dado intervalo.

Por exemplo, se na média encontramos 2 pedaços de alface por bandeja no bandeco, um experimento de Poisson pode ser ver quantos pedaços de alface há em uma certa bandeja.

* Diferença da Binomial: o exemplo anterior ilustra a diferença entre Poisson e Binomial. A binomial vê sucessos (tem ou não tem alface); já a Poisson se relaciona a contagem de ocorrências por evento (não se tem, mas quanto tem)

Se X tem distribuição de Poisson com taxa λ ($X = \text{Poisson}(\lambda)$), temos que:

$$\boxed{\begin{aligned} P(X=k) &= \frac{e^{-\lambda}\lambda^{k}}{k!} \\ E(X) &= \lambda \\ Var(X) &= \lambda \end{aligned}}$$

Além disso, se tomarmos $X_1 = \text{Poisson}(\lambda_1), \ldots, X_m = \text{Poisson}(\lambda$ independentes, a soma delas também será Poisson e:

$$\sum_{i=1}^{m} X_i = \text{Poisson}\left(\sum_{i=1}^{m} \lambda_i\right)$$

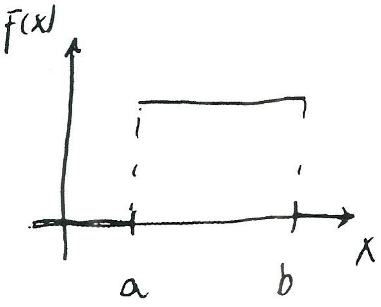
IV - Distribuição de Probabilidade Contínuas

Vejamos algumas distribuições de probabilidade muito comuns para v.a. contínuas.

1) Distribuição Uniforme

Seja X tal que sua distribuição de probabilidade seja:

$$F(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & \text{caso contrário} \end{cases}$$



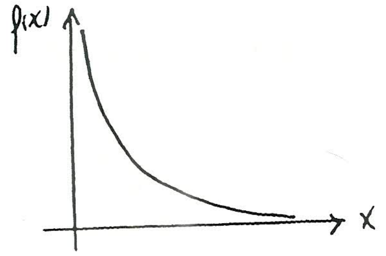
Pode-se provar que:

$$\boxed{\begin{aligned} E(X) &= \frac{a+b}{2} \\ Var(X) &= \frac{(b-a)^2}{12} \\ F(X) &= \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases} \end{aligned}}$$

2) Distribuição Exponencial

Dizemos que uma v.a. contínua tem distribuição exponencial com parâmetro λ se sua Função densidade de probabilidade for da forma:

$$F(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , \text{caso contrário} \end{cases}$$



Pode-se provar que:

$$\boxed{\begin{aligned} E(X) &= \frac{1}{\lambda} \\ Var(X) &= \frac{1}{\lambda^2} \\ F(x) &= \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , \text{caso contrário} \end{cases} \end{aligned}}$$

3) Distribuição Normal

Dizemos que uma v.a. contínua tem distribuição normal com parâmetros μ (média) e σ (desvio padrão) se sua f.d.p é dada por:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R}$$

Nesse caso, representamos:

$$X = N(\mu; \sigma^2)$$

Tinha calma, não teremos que usar essa fórmula, pois temos tabelados os valores para a normal padrão ($N(0,1)$) e a partir deles podemos achar outros como veremos mais adiante.

Temos em tabelas os valores de $P(0 \leq X \leq a)$ para a normal padrão $N(0,1)$. Para outros valores dos parâmetros fazemos:

$$\boxed{P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)}$$

Em que $\frac{X-\mu}{\sigma} = Z$ segue a normal padrão.

* Soma de normais: Se $X_1, \dots, X_n$ são v.a. independentes com distribuição normal: $X_i = N(\mu_i, \sigma_i^2)$ temos que $\sum_{i=1}^{n} a_i X_i$ tem distribuição normal com parâmetros

$$\mu = \sum_{i=1}^{n} a_i \mu_i \qquad \sigma^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$$

V - Teorema Central do Limite

Por muitos motivos, na estatística gosta-se de trabalhar com distribuições normais. Felizmente, a normal é mais comum do que se imagina. Um exemplo em que a normal aparece para facilitar nossas vidas é dado pelo Teorema Central do Limite (TCL)

Segundo o TCL, se $X_1, X_2, \ldots, X_n$ são variáveis aleatórias independentes com média μ e variância σ^2, a média de X_i tenderá a uma distribuição normal na forma,

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{\text{tende a}} N\left(\mu ; \frac{\sigma^2}{n}\right)$$

Aproximação da Binomial por Normal

Se $X_1, X_2, \ldots, X_n$ forem resultados de experimentos de Bernoulli com probabilidade de sucesso p, já vimos que a soma dos X_s terá uma distribuição binomial com média np e variância $np(1-p)$

Se aplicarmos o TCL, sabemos que:

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \text{ tende à normal } N\left(\mu, \frac{\sigma^2}{n}\right)$$

Logo para a soma (resultado da binomial):

$$X_1 + X_2 + \ldots + X_n \text{ tende à normal } N(n\mu ; n\sigma^2)$$

Como para cada X_i de Bernoulli $\mu = p$ e $\sigma^2 = p(1-p)$, segue que a binomial se aproxima da normal $\boxed{N(np ; np(1-p))}$

* Quando usar: o TCL diz que a variável X binomial tende a uma distribuição normal. Quando, porém, podemos tratar de fato X como uma normal?

Quando n for suficientemente grande. Não há um valor exato para dizer se o n é grande ou pequeno, mas algumas regrinhas podem ser usadas

i) $np > 10$ e $np(1-p) > 10$; ou

ii) $n > 30$

* Por que o TCL é legal? imagine que X é um experimento binomial com 100 repetições ($n = 100$) e que queremos calcular a probabilidade de X ser menor do que 50

Ou seja: $P(X < 50) = F(50)$

Se não usássemos o TCL, deveríamos somar $P(x_j)$ para x_j entre 0 e 49. Não parece divertido, não é mesmo... Por outro lado, se sei que X pode ser aproximada por uma normal, bastará consultar a tabela da acumulada.

VI - Distribuições de Probabilidade conjunta

Até agora estudamos distribuições de probabilidade com uma única variável aleatória. Entretanto, muitas vezes o fenômeno que queremos estudar depende de um número maior de variáveis aleatórias.

Neste tópico vamos estudar as variáveis aleatórias bidimensionais, que dependem de duas variáveis aleatória (X e Y, digamos)

1) Distribuição Conjunta (Caso DISCRETO)

Para variáveis aleatórias, vimos que a função densidade de probabilidade $p(X)$ nos fornecia a probabilidade da variável X assumir cada valor no domínio.

Para o caso de duas variáveis aleatórias discretas, a função de probabilidade conjunta $p(x,y)$ é definida como a probabilidade de:

$$\boxed{p(x,y) = P(X = x \text{ e } Y = y)}$$

Analogamente ao que vimos com uma v.a., a soma de todas as probabilidades deve ser 1:

$$\sum_{x_i, y_j} P(X = x_i, Y = y_j) = 1$$

2) Distribuições Marginais (caso DISCRETO)

Apesar da distribuição depender de X e Y, muitas vezes vamos querer a probabilidade de X assumir certo valor, independentemente do valor de Y, e vice-versa.

Para isso, podemos usar as funções massa de probabilidade marginais de X e Y, representadas por $p_X(x)$ e $p_Y(y)$ e calculadas como:

$$\boxed{p_X(x) = \sum_y p(x,y) \qquad p_Y(y) = \sum_x p(x,y)}$$

3) Independência:

Dizemos que duas variáveis aleatórias X e Y são independentes se:

$$\boxed{p(x,y) = p_X(x) * p_Y(y)}$$

Exemplo: Fonte: DEVORE

Considere a seguinte função massa de probabilidade conjunta de X e Y

X \ Y	0	1	2
0	0,10	0,04	0,02
1	0,08	0,20	0,06
2	0,06	0,14	0,30

a) Qual é $P(X=1 \text{ e } Y=1)$

b) Calcule $P(X \leq 1 \text{ e } Y \leq 1)$

c) Calcule a Fmp marginal de X e de Y

d) Calcule $P(X \leq 1)$

e) X e Y são variáveis independentes?

Solução

a) Da tabela: $p(1,1) = P(X=1 \text{ e } Y=1) = 0{,}2$

b) $P(X \leq 1 \text{ e } Y \leq 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1)$

$P(X \leq 1 \text{ e } Y \leq 1) = 0{,}1 + 0{,}04 + 0{,}08 + 0{,}2$

$P(X \leq 1 \text{ e } Y \leq 1) = 0{,}42$

c) Para calcular a marginal, basicamente fixamos o valor de uma variável e somamos as probabilidades para todos os valores da outra.

Na tabela, fazemos isso somando linhas ou colunas:

X \ Y	0	1	2	$P_X(x)$
0	0,10	0,04	0,02	0,16
1	0,08	0,20	0,06	0,34
2	0,06	0,14	0,30	0,5
$P_Y(y)$	0,24	0,38	0,38	

← $P_X(0)$, ou seja, X é zero e y é qualquer

$P_Y(1)$, ou seja, $Y=1$ e X qualquer

d) $P(X \leq 1) = P(X=0) + P(X=1) = P_X(0) + P_X(1) = 0{,}5$

e) Não, note por exemplo que $p(0,0) \neq P_X(0) * P_Y(0)$

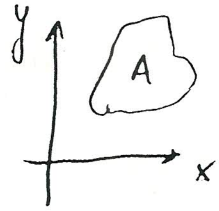
$0{,}1 \neq 0{,}16 * 0{,}24$

Agora vamos estudar estudar o caso de variáveis aleatórias bidimensionais <u>contínuas</u>

4) <u>Distribuição Conjunta</u> (caso CONTÍNUO)

Se X e Y são v.a. contínuas, $f(x,y)$ é a função de densidade de probabilidade conjunta. Para qualquer conjunto A de valores de (x,y) temos que:

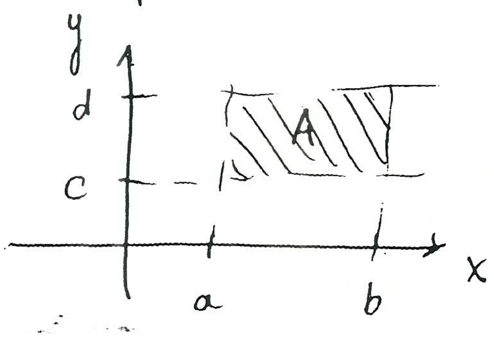
$$\boxed{P[(X,Y) \in A] = \iint_A f(x,y)\,dx\,dy}$$



No caso especial de A ser um retângulo com $a \leq x \leq b$ e $c \leq y \leq d$, temos:

$$P[(X,Y) \in A] = \int_a^b \int_c^d f(x,y)\,dy\,dx$$

Em que A é o retângulo:



<u>Propriedades</u>

i) $f(x,y) \geq 0 \quad \forall\, x,y \in \mathbb{R}$

ii) $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

5) Distribuição Marginal (caso CONTÍNUO)

Para X e Y vas contínuas, as funções densidade de probabilidade marginais são dadas por:

$$\boxed{f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx}$$

fixo x e integro (somo) em y

Fixo y e integro em x

Exemplo (Fonte: DEVORE)

Considere a seguinte distribuição de probabilidade:

$$f(x,y) = \begin{cases} K(x^2+y^2), & 20 \leq x \leq 30,\ 20 \leq y \leq 30 \\ 0, & \text{caso contrário} \end{cases}$$

a) Qual o valor de K?

b) Calcule $P(X \leq 26 \text{ e } Y \leq 26)$

c) Calcule $P(|X-Y| \leq 2)$

d) Determine $f_X(x)$ e $f_Y(y)$

e) X e Y são independentes?

a) Aplicamos que:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

$$\int_{20}^{30}\int_{20}^{30} K(x^2+y^2)\,dx\,dy = 1$$

$$\int_{20}^{30} K\left(\frac{x^3}{3} + xy^2\right)\Big|_{20}^{30} dy = 1$$

$$\int_{20}^{30} K\left(\frac{19000}{3} + 10y^2\right) dy = 1$$

$$K\left(\frac{19000}{3}y + \frac{10y^3}{3}\right)\Big|_{20}^{30} = 1$$

$$K\left(\frac{190000}{3} + \frac{190000}{3}\right) = 1$$

$$\boxed{K = 3/380000}$$

b) $P(X \leq 26 \text{ e } Y \leq 26) = \int_{-\infty}^{26}\int_{-\infty}^{26} f(x,y)\,dx\,dy$

$$P(X \leq 26 \text{ e } Y \leq 26) = \int_{0}^{26}\int_{0}^{26} \frac{3}{380000}(x^2+y^2)\,dx\,dy$$

(...)

$$P(X \leq 26 \text{ e } Y \leq 26) = \frac{189}{625} = 0,3024$$

c)

Queremos calcular a probabilidade de:

$$|X - Y| \leq 2$$

$$-2 \leq X - Y \leq 2$$

$$-2 + Y \leq X \leq 2 + Y$$

Ou seja:

$$P(|X-Y| \leq 2) = \int_{20}^{30} \int_{y-2}^{y+2} K(x^2+y^2)\,dx\,dy$$

$$P(|X-y| \leq 2) = \int_{20}^{30} K\left(\frac{x^3}{3} + xy^2\right)\Big|_{y-2}^{y+2} dy$$

$$P(|X-y| \leq 2) = K\int_{20}^{30} \left(\frac{(y+2)^3}{3} + (y^3+2y^2) - \frac{(y-2)^2}{3} - (y^3-2y^2)\right) dy$$

$$P(|X-y| \leq 2) = 0{,}3593$$

d/e) $$F_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_{20}^{30} K(x^2+y^2)\,dy = K\left(x^2y + \frac{y^3}{3}\right)\Big|_{20}^{30}$$

$$F_X(x) = K\left(30x^2 + \frac{30^3}{3} - 20x^2 - \frac{20^3}{3}\right)$$

$$\boxed{F_X(x) = 10Kx^2 + 0{,}05}$$

Analogamente:

$$\boxed{F_Y(y) = 10Ky^2 + 0{,}05}$$

$$F_X(x) * F_Y(y) = (10Kx^2 - 0{,}05) * (10Ky^2 - 1$$

$$F_X(x) * F_Y(y) = -0{,}5Kx^2 - 0{,}5Ky^2 + \ldots + 100K^2x^2y^2 + 0{,}0025$$

Claramente $F_X(x) * F_Y(y) \neq f(x,y$

Logo X e Y não são independentes