Revisão de Probabilidades e Estatística

Afshine Amidi e Shervine Amidi

13 de Outubro de 2018

Introdução a Probabilidade e Combinatória

❒ **Espaço amostral** – O conjunto de todos os resultados possíveis é chamado de espaço amostral do experimento e é denotado por S.

❒ **Evento** – Qualquer subconjunto E do espaço amostral é chamado de evento. Isso é, um evento é um conjunto de possíveis resultados do experimento. Se o resultado do experimento está contido em E, então é dito que o evento ocorreu.

❒ **Axiomas de probabilidade** – Para cada evento E, denotamos $P(E)$ a probabilidade do evento E ocorrer.

$$(1)\quad \boxed{0\leqslant P(E)\leqslant 1}\qquad (2)\quad \boxed{P(S)=1}\qquad (3)\quad \boxed{P\left(\bigcup_{i=1}^{n}E_i\right)=\sum_{i=1}^{n}P(E_i)}$$

❒ **Permutação** – A permutação é um arranjo de r objetos de um conjunto de n objetos, em uma determinada ordem. O número desses arranjos é dado por $P(n,r)$, definido como:

$$\boxed{P(n,r)=\frac{n!}{(n-r)!}}$$

❒ **Combinação** – A combinação de um arranjo de r objetos de um conjunto de n objetos, onde a ordem não importa. O número desses arranjos é dado por $C(n,r)$, definido como:

$$\boxed{C(n,r)=\frac{P(n,r)}{r!}=\frac{n!}{r!(n-r)!}}$$

Observação: dado que $0\leqslant r\leqslant n$, então temos que $P(n,r)\geqslant C(n,r)$.

Probabilidade Condicional

❒ **Regra de Bayes** – Para eventos A e B tal que $P(B)>0$, temos que:

$$\boxed{P(A|B)=\frac{P(B|A)P(A)}{P(B)}}$$

Observação: temos que $P(A\cap B)=P(A)P(B|A)=P(A|B)P(B)$.

❒ **Partição** – Dado que $\{A_i,i\in[\![1,n]\!]\}$ seja tal que para todo i, $A_i\neq\varnothing$. Dizemos que $\{A_i\}$ é uma partição se temos:

$$\boxed{\forall i\neq j, A_i\cap A_j=\emptyset\quad \textrm{e}\quad \bigcup_{i=1}^{n}A_i=S}$$

Observação: para qualquer evento B no espaço amostral temos que $P(B)=\sum_{i=1}^{n}P(B|A_i)P(A_i)$.

❒ **Extensão da regra de Bayes** – Seja $\{A_i,i\in[\![1,n]\!]\}$ uma partição do espaço amostral. Temos que:

$$\boxed{P(A_k|B)=\frac{P(B|A_k)P(A_k)}{\displaystyle\sum_{i=1}^{n}P(B|A_i)P(A_i)}}$$

❒ **Independência** – Dois eventos A e B são independentes se e apenas se tivermos:

$$\boxed{P(A\cap B)=P(A)P(B)}$$

Variável aleatória

❒ **Variável aleatória** – Uma variável aleatória, normalmente denominada X, é uma função que mapeia todo elemento em um espaço amostral para uma linha verdadeira.

❒ **Função de distribuição cumulativa (CDF)** – A função de distribuição cumulativa F, que é monotonicamente não decrescente e é tal que

$$\lim_{x\rightarrow-\infty}F(x)=0\quad\textrm{e}\quad\lim_{x\rightarrow+\infty}F(x)=1$$

é definida como:

$$\boxed{F(x)=P(X\leqslant x)}$$

Lembrete: temos que $P(a<X\leqslant B)=F(b)-F(a)$.

❒ **Função densidade de probabilidade (PDF)** – A função densidade de probabilidade f é a probabilidade de que X assuma valores entre duas realizações adjacentes da variável aleatória.

❒ **Relações envolvendo a PDF e a CDF** – Aqui estão as propriedades mais importantes que se deve conhecer dos casos discretos (D) e contínuos (C).

Caso	CDF F	PDF f	Propriedades da PDF
(D)	$F(x)=\sum_{x_i\leqslant x}P(X=x_i)$	$f(x_j)=P(X=x_j)$	$0\leqslant f(x_j)\leqslant 1$ e $\sum_j f(x_j)=1$
(C)	$F(x)=\int_{-\infty}^{x}f(y)dy$	$f(x)=\frac{dF}{dx}$	$f(x)\geqslant 0$ e $\int_{-\infty}^{+\infty}f(x)dx=1$

❒ **Variância** – A variância de uma variável aleatória, normalmente denominada $\textrm{Var}(X)$ ou σ^2, é a medida do espalhamento da sua função de distribuição. Ela é determinada por:

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

❒ **Desvio padrão** – O desvio padrão de uma variável aleatória, normalmente denominado σ, é a medida do espalhamento da sua função de distribuição que é compatível com a unidade da variável aleatória. Ele é determinado por:

$$\sigma=\sqrt{\textrm{Var}(X)}$$

❒ **Expectativas e Momentos da Distribuição** – Aqui estão as expressões do valor esperado $E[X]$, do valor esperado generalizado $E[g(X)]$, do k-ésimo momento $E[X^k]$ e função característica $\psi(\omega)$ para os casos discretos e contínuos:

Caso	$E[X]$	$E[g(X)]$	$E[X^k]$	$\psi(\omega)$
(D)	$\displaystyle\sum_{i=1}^nx_if(x_i)$	$\displaystyle\sum_{i=1}^ng(x_i)f(x_i)$	$\displaystyle\sum_{i=1}^nx_i^kf(x_i)$	$\displaystyle\sum_{i=1}^nf(x_i)e^{i\omega x_i}$
(C)	$\displaystyle\int_{-\infty}^{+\infty}xf(x)dx$	$\displaystyle\int_{-\infty}^{+\infty}g(x)f(x)dx$	$\displaystyle\int_{-\infty}^{+\infty}x^kf(x)dx$	$\displaystyle\int_{-\infty}^{+\infty}f(x)e^{i\omega x}dx$

Remarque: on a $e^{i\omega x}=\cos(\omega x)+i\sin(\omega x)$.

❒ **Transformação das variáveis aleatórias** – Sejam as variáveis X e Y ligadas por alguma função. Ao denotador f_X e f_Y para as funções de distribuição de X e de Y respectivamente, temos que:

$$f_Y(y)=f_X(x)\left|\frac{dx}{dy}\right|$$

❒ **Regra integral de Leibniz** – Seja g uma função de x e possivelmente de c, e a,b fronteiras que podem depender de c. Temos que:

$$\frac{\partial}{\partial c}\left(\int_a^bg(x)dx\right)=\frac{\partial b}{\partial c}\cdot g(b)-\frac{\partial a}{\partial c}\cdot g(a)+\int_a^b\frac{\partial g}{\partial c}(x)dx$$

❒ **Desigualdade de Chebyshev** – Seja X uma variável aleatória com valor esperado μ. Para $k,\sigma>0$, temos a seguinte desigualdade:

$$P(|X-\mu|\geqslant k\sigma)\leqslant\frac{1}{k^2}$$

Variáveis aleatórias distribuídas conjuntamente

❒ **Densidade condicional** – A densidade condicional de X com respeito a Y, normalmente denotada como $f_{X|Y}$, é definida como:

$$f_{X|Y}(x)=\frac{f_{XY}(x,y)}{f_Y(y)}$$

❒ **Independência** – Duas variáveis aleatórias X e Y são ditas independentes se:

$$f_{XY}(x,y)=f_X(x)f_Y(y)$$

❒ **Densidade marginal e distribuição cumulativa** – A partir da função de probabilidade de densidade conjunta f_{XY}, temos que:

Caso	**Densidade marginal**	**Função cumulativa**
(D)	$\displaystyle f_X(x_i)=\sum_jf_{XY}(x_i,y_j)$	$\displaystyle F_{XY}(x,y)=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y}f_{XY}(x_i,y_j)$
(C)	$\displaystyle f_X(x)=\int_{-\infty}^{+\infty}f_{XY}(x,y)dy$	$\displaystyle F_{XY}(x,y)=\int_{-\infty}^x\int_{-\infty}^yf_{XY}(x',y')dx'dy'$

❒ **Coveriância** – Definimos covariância de duas variáveis aleatórias X e Y, que chamamos de σ_{XY}^2 ou mais comumente de $\textrm{Cov}(X,Y)$, como:

$$\textrm{Cov}(X,Y)\triangleq\sigma_{XY}^2=E[(X-\mu_X)(Y-\mu_Y)]=E[XY]-\mu_X\mu_Y$$

❒ **Correlação** – Dado que σ_X,σ_Y são os desvios padrão de X e Y, definimos a correlação entre as variáveis aleatórias X e Y, denominada ρ_{XY}, como:

$$\rho_{XY}=\frac{\sigma_{XY}^2}{\sigma_X\sigma_Y}$$

Observação 1: é definido que para qualquer variáveis aleatórias X, Y temos que $\rho_{XY}\in[-1,1]$.
Observação 2: Se X e Y são independentes, então $\rho_{XY}=0$.

❒ **Distribuições principais** – Aqui estão as principais distribuições que não devem ser esquecidas:

Tipo	**Distribuição**	**PDF**	$\psi(\omega)$	$E[X]$	$\textrm{Var}(X)$
(D)	$X\sim\mathcal{B}(n,p)$ Binomial	$\displaystyle P(X=x)=\binom{n}{x}p^xq^{n-x}$ $x\in[\![0,n]\!]$	$(pe^{i\omega}+q)^n$	np	npq
	$X\sim\textrm{Po}(\mu)$ Poisson	$\displaystyle P(X=x)=\frac{\mu^x}{x!}e^{-\mu}$ $x\in\mathbb{N}$	$e^{\mu(e^{i\omega}-1)}$	μ	μ
(C)	$X\sim\mathcal{U}(a,b)$ Uniform	$\displaystyle f(x)=\frac{1}{b-a}$ $x\in[a,b]$	$\displaystyle\frac{e^{i\omega b}-e^{i\omega a}}{(b-a)i\omega}$	$\displaystyle\frac{a+b}{2}$	$\displaystyle\frac{(b-a)^2}{12}$
	$X\sim\mathcal{N}(\mu,\sigma)$ Gaussian	$\displaystyle f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ $x\in\mathbb{R}$	$e^{i\omega\mu-\frac{1}{2}\omega^2\sigma^2}$	μ	σ^2
	$X\sim\textrm{Exp}(\lambda)$ Exponential	$f(x)=\lambda e^{-\lambda x}$ $x\in\mathbb{R}_+$	$\displaystyle\frac{1}{1-\frac{i\omega}{\lambda}}$	$\displaystyle\frac{1}{\lambda}$	$\displaystyle\frac{1}{\lambda^2}$

Estimativa de parâmetro

❒ **Amostra aleatória** – Uma amostra aleatória é uma coleção de n variáveis aleatórias $X_1, ..., X_n$ que são independentes e igualmente distribuidas com X.

❒ **Estimador** – Um estimador é uma função dos dados que é usada para inferir o valor de um parâmetro desconhecido em um modelo estatístico.

❒ **Viés** – O viés de um estimador $\hat{\theta}$ é definido como a diferença entre o valor esperado da distribuição de $\hat{\theta}$ e o seu real valor, i.e.:

$$\boxed{\textrm{Bias}(\hat{\theta}) = E[\hat{\theta}] - \theta}$$

Observação: um estimador é chamado de imparcial (unbiased) quando $E[\hat{\theta}] = \theta$.

❒ **Média da amostra** – A média da amostra de uma amostra aleatória é usada para estimar a verdadeira média μ de uma distribuição, e é denominada $\overline{X}$ e é definida como:
Observação: a média da amostra é imparcial, i.e $E[\overline{X}] = \mu$.

$$\boxed{\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i}$$

❒ **Amostra da variância** – A amostra da variância de uma amostra aleatória é usada para estimar a verdadeira variância σ^2 da distribuição, e é normalmente denominada s^2 ou $\hat{\sigma}^2$ e definida por:
Observação: a variância da amostra é imparcial, i.e $E[s^2] = \sigma^2$.

$$\boxed{s^2 = \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2}$$

❒ **Teorema do Limite Central** – Dado que temos uma amostra aleatória $X_1, ..., X_n$ seguindo uma determinada distribuição com a média μ e a variância σ^2, temos que:

$$\boxed{\overline{X} \underset{n \to +\infty}{\sim} \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)}$$