UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

MAT02214 - Estatística Geral 1 - 2022/2

Plano Aula 25 e 26

(cont... Variáveis Aleatórias)

Vetores aleatórios (capítulo 8, Livro Bussab e Morettin)

Estudamos agora o comportamento de mais de uma variável aleatória, $X, Y, Z, \ldots$.

1. O caso de duas variáveis X e Y discretas

Definição **vetor aleatório discreto**: denotamos (X, Y) um vetor aleatório discreto, em que X e Y são v.a. discretas definidas no mesmo espaço amostral S de um experimento aleatório E, assumindo particulares pares de valores (x, y).

Exemplo 1: Em uma pesquisa de opinião com *toda uma população*, estamos interessados na opinião quanto ao desempenho do atual governo, nas categorias péssimo, ruim, regular, bom e ótimo ($Y \in \{1, 2, 3, 4, 5\}$), em relação ao fato de o(a) respondente não ter cursado escola pública ou ter cursado ($X \in \{0, 1\}$). Ou quanto aos anos de estudo ($Z \in \{1, 2, \ldots\}$).
* Será que a opinião quanto ao desempenho do atual governo (Y) é o mesmo para os que não estudaram em escola pública ou estudaram (X)? E em relação a escolaridade (Z)?
* A relação entre X e Y tende a ser direta ou indireta (inversa)? E entre Z e Y?

a. Distribuição conjunta (seção 8.1, Livro Bussab e Morettin)

Definição **Função massa de probabilidade conjunta**: Seja (X, Y) um vetor aleatório discreto, a f.m.p. conjunta $p(x, y)$ é definida para cada par (x, y) por

$$p(x, y) = P([X = x] \cap [Y = y]).$$

- *propriedades*: (1) $0 \leq p(x, y) \leq 1$; (2) $\sum_x \sum_y p(x, y) = 1$.

b. Distribuições Marginais (seção 8.2, Livro Bussab e Morettin)

Definição **Funções massa de probabilidade marginais**: Seja (X, Y) um vetor aleatório discreto, as f.m.p. marginais de X e Y são dadas respectivamente por

$$p(x) = \sum_y p(x, y) \text{ e } p(y) = \sum_x p(x, y).$$

- $p(x)$ e $p(y)$ seguem as mesmas propriedades de uma f.m.p.

2. O caso de duas variáveis X e Y contínuas (seção 8.5, Livro Bussab e Morettin)

Definição **vetor aleatório contínuo**: denotamos (X, Y) um vetor aleatório contínuo, onde X e Y são v.a. contínuas definidas no mesmo espaço amostral S de um experimento aleatório E, assumindo particulares pares de valores $\{(x, y) : a \leq x \leq b, c \leq y \leq d\}$.

Exemplo 2: Após a adoção do autoatendimento em uma rede de lanchonetes, os administradores estão interessados em estudar a relação entre a proporção de tempo em que o caixa está ocupado ao longo do dia (Y) com a proporção de tempo em que o terminal de autoatendimento está em uso (X).
* A proporção de tempo em que o terminal de autoatendimento está em uso (X) pode estar relacionada com a proporção de tempo que o caixa está atendendo (Y)?
* A relação entre Z e W tende a ser direta ou indireta (inversa)? Linear?

a. Distribuição conjunta (pág. 225, seção 8.5, Livro Bussab e Morettin)

Definição **Função densidade de probabilidade conjunta**: Seja (X, Y) um vetor aleatório contínuo, denominamos $f(x, y)$ a f.d.p. conjunta de X e Y se

$$P(a \leq x \leq b, c \leq y \leq d) = \int_a^b \int_c^d f(x, y)dxdy.$$

- *propriedades*: (1) $f(x, y) > 0$; (2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dxdy = 1$.

b. Distribuições marginais (pág. 227, seção 8.5, Livro Bussab e Morettin)

Definição **Funções densidade de probabilidade marginais**: Seja (X, Y) um vetor aleatório contínuo, as f.d.p. marginais de X e Y são dadas respectivamente por

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

- $f(x)$ e $f(y)$ seguem as mesmas propriedades de uma f.d.p.

3. Variáveis aleatória independentes

Sabemos que dois conjuntos A e B, definidos no mesmo espaço amostral S de um experimento aleatório E, são independentes se e somente se $P(A \cap B) = P(A) \times P(B)$. Para v.a. temos

Definição **Variáveis aleatórias independentes**: Duas v.a. são independentes se para cada par de valores x e y, temos

$$p(x, y) = p(x) \times p(y), \text{ se } (X, Y) \text{ é um vetor aleatório discreto,}$$

ou

$$f(x, y) = f(x) \times f(y), \text{ se } (X, Y) \text{ é um vetor aleatório contínuo.}$$

4. Covariância e correlação (seção 8.4, Livro Bussab e Morettin)

- **Funções de variáveis aleatórias** (seção 8.3, Livro Bussab e Morettin)
  - Se temos interesse na v.a. $Z = X + Y$, como calcular $E(Z) = E(X + Y)$ e $Var(Z) = Var(X + Y)$?
  - ou $W = XY$, qual a $E(W) = E(XY)$ e $Var(XY)$?
- Como medir a dependência entre X e Y?

Covariância

Definição **covariância**: Sejam X e Y duas v.a., se $E(X)$, $E(Y)$, $V(X)$ e $V(Y)$ existem e são finitas, então a covariância entre X e Y é dada por

$$Cov(X, Y) = E\left[(X - EX)(Y - EY)\right]$$

, ou alternativamente

$$Cov(X, Y) = E\left[XY\right] - E(X)E(Y).$$

- *Proposições*:
  - $-\infty \leq Cov(X, Y) \leq \infty$;
  - $Cov(X, X) = E\left[(X - EX)(x - Ex)\right] = E\left[(X - EX)^2\right] = Var(X)$;
  - $Cov(aX, bY) = abCov(X, Y)$, para a e b constantes;
  - $Cov(X, a) = 0$;
  - $Cov(X + Y) = Cov(X - Y) = Var(X) + Var(Y) - 2Cov(X, Y)$;
  - se X e Y são *independentes*, $Cov(X, Y) = 0$ pois $E(XY) = E(X)E(Y)$;

- também se X e Y são *independentes*, $Var(X+Y) = Var(X-Y) = Var(X) + Var(Y)$;
- $Cov(X+Y, Z) = Cov(X, Z) + Cov(Y, Z)$, para Z outra v.a..

ATENÇÃO!!! Se $Cov(X, Y) = 0$ não significa que X e Y são independetes, dizemos que X e Y são não correlacionados.

Correlação

Definição **correlação**: Sejam X e Y duas v.a., se $E(X)$, $E(Y)$, $V(X)$ e $V(Y)$ existem e são finitas, então o coeficiente de correlação entre X e Y é dado por

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

- *Proposições*:
  - O coeficiente de correlação é *adimensional* $-1 \leq Cov(X, Y) \leq 1$;
  - se X e Y são *independentes* $\rho(X, Y) = 0$, pois $Cov(X, Y) = 0$;
  - se $Y = aX + b$, para a e b constantes e $a \neq 0$, dizemos que existe correlação perfeita entre X e Y, então teremos $\rho(X, Y) = -1$ ou $\rho(X, Y) = 1$.
- Observações:
  - $\rho(X, Y)$ indica o grau de linearidade da relação entre X e Y. Um $\rho(X, Y)$ próximo de 1 ou -1 indica um forte grau de linearidade;
  - $\rho(X, Y) > 0$ indica uma dependência linear direta (positiva), tendência de aumento de uma v.a. à medida que a outra aumenta também; $\rho(X, Y) < 0$ indica uma dependência linear inversa (negativa), tendência de decrescimento de uma v.a. à medida que a outra aumenta;
  - $\rho(X, Y) \approx 0$ não indica necessariamente ausência de dependência entre X e Y, mas se essa relação existir não deve ser linear.

Ler apostila "Notas de Aula MAT02214 - Estatística Geral I" capítulo 5 seção 1.

Continuar lista de exercícios 2-4.
