U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Capítulo 4

Variáveis Aleatórias. Distribuições de Probabilidade

AMG, JFO (v8 – 2017)
adaptado de: *Estatística*,
Rui Campos Guimarães,
José A. Sarsfield Cabral

Slide 5.-1

Conteúdo

Slide 5.0

Resultados de aprendizagem

- Definir v.a.(s) a partir da descrição de uma experiência aleatória
- Determinar funções (densidade) de probabilidade de v.a.
- Determinar funções de distribuição a partir de funções (densidade) de probabilidade e *vice-versa*
- Calcular probabilidades a partir de funções de probabilidade, densidade de probabilidade e de distribuição
- Recordar os conceitos de valor esperado, variância e desvio padrão de v.a.
- Calcular valores esperados, variâncias e desvios padrões de v.a. a partir das respectivas funções (densidade) de probabilidade ou a partir das propriedades do valor esperado e da variância
- Demonstrar propriedades do valor esperado e da variância
- Dada a descrição de uma experiência aleatória definir a(s) v.a.(s) bidimnsionais associada(s)
- Determinar funções de distribuição conjuntas a partir de funções (de densidade) de probabilidade conjuntas e *vice-versa*
- Determinar funções marginais e condicionais (de densidade) de probabilidade a partir de funções (de densidade) de probabilidade conjuntas

Slide 5.1

- Explicar o conceito de probabilidades condicionais aplicado a funções condicionais (de densidade) de probabilidade
- Calcular probabilidades a partir de funções (de densidade) de probabilidade conjuntas e de distribuição conjuntas, assim como de funções condicionais (de densidade) de probabilidade
- Explicar o conceito de independência entre variáveis aleatórias
- Verificar a independência entre variáveis aleatórias
- Calcular valores esperados, variâncias e desvios padrões de v.a. a partir das respectivas funções (de densidade) de probabilidade conjuntas ou das propriedades do valor esperado e da variância
- Calcular covariâncias e coeficientes de correlação entre pares de variáveis aleatórias
- Demonstrar propriedades do valor esperado, da variância e da covariância envolvendo pares de variáveis aleatórias

Slide 5.2

- Explicar o conceito de variável transformada
- Calcular valores esperados e variâncias de variáveis obtidas por transformações lineares conhecendo apenas os parâmetros da(s) v.a.(s) original(is)
- ~~Determinar funções (densidade) de probabilidade de funções de v.a. (variáveis transformadas)~~
- ~~Calcular de forma aproximada valores esperados e variâncias de variáveis obtidas por transformações não-lineares conhecendo apenas os parâmetros da v.a. original~~
- ~~Determinar funções (de densidade) de probabilidade de variáveis transformadas (apenas casos simples)~~
- ~~Calcular de forma aproximada valores esperados e variâncias de variáveis obtidas por transformações não-lineares conhecendo apenas os parâmetros das v.a.(s) originais~~

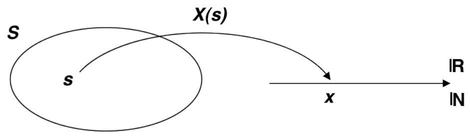
Nota:

- Neste capítulo serão estudadas em maior detalhe as v.a. discretas

Slide 5.3

4.1 Variável Aleatória

- Permite exprimir os resultados de uma experiência aleatória
- Aplicação do espaço amostral sobre um conjunto de chegada



Notação:
• X – variável aleatória (v.a.)
• x – valor que a v.a. toma

$$\begin{array}{c}\text{Contradomínio}\\ \text{da aplicação}\end{array} = \begin{array}{c}\text{domínio}\\ \text{da v.a.}\end{array} \begin{cases} \text{qualitativo} \longrightarrow \textit{v.a. qualitativa} \\ \text{discreto} \longrightarrow \textit{v.a. discreta} \\ \text{contínuo} \longrightarrow \textit{v.a. contínua} \end{cases}$$

- Sobre um mesmo espaço amostral podem ser definidas diferentes aplicações, ou seja diferentes variáveis aleatórias
- Os valores que a v.a. toma podem ou não confundir-se com o Espaço Amostral

Slide 5.4

Variável Aleatória – Exemplos

Lançamento de uma moeda E-C ao ar 3 vezes. Espaço amostral S: sequências de E's e C's.

Y: número de E's em cada sequência ($Y : \{0,1,2,3\}$)

A cada valor de Y associa-se uma probabilidade, calculada a partir das probabilidades dos elementos do espaço amostral:

$P(Y=0) = P(CCC) = 1/8$

$P(Y=1) = P(ECC) + P(CEC) + P(CCE) = 3/8$

$P(Y=2) = P(EEC) + P(ECE) + P(CEE) = 3/8$

$P(Y=3) = P(EEE) = 1/8$

Considere novamente a experiência aleatória de lançar ao ar 3 vezes uma moeda E-C e anotar a sequência de E's e C's.

Z: número de E's consecutivos em cada sequência ($Z : \{0,1,2,3\}$)

Calcule as probabilidades associadas a cada valor de Z

Slide 5.5

V.A. qualitativa

Experiência Aleatória – Nascimento de uma criança

Espaço Amostral – Feminino ou Masculino

Variável Aleatória – Género

V.A. discreta

Experiência Aleatória – Lançamento da moeda E-C três vezes ao ar

Espaço Amostral – Sequências de E's e C's

Variável Aleatória – Número de E's obtido em cada sequência

V.A. contínua

Experiência Aleatória – Medição da altura de uma pessoa

Espaço Amostral – Conjunto de todas as alturas ($[Alt_{min}; Alt_{max}]$)

Variável Aleatória – Altura

Slide 5.6

4.2 Distribuições Discretas

4.2.1 Função de probabilidade e função de distribuição

V.A. discretas – Função de probabilidade

Função de probabilidade ($p(y)$)
Associa a cada valor que a variável aleatória pode tomar (y), a probabilidade de Y ser igual a y (em que$Y \in \{y_1, y_2, y_3, \ldots\}$)

$$p(y) = \text{Probabilidade}(Y = y) = P(Y = y)$$

Propriedades (resultam dos axiomas de probabilidade)

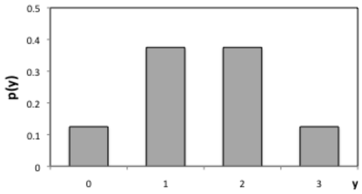
(1) $p(y_i) \geq 0, \, \forall_i$

(2) $\sum_{i=1}^{n} p(y_i) = 1$

Exemplo – Lançamento da moeda E-C três vezes ao ar

Y: número de E's em cada sequência

y	0	1	2	3
$p(y)$	$1/8$	$3/8$	$3/8$	$1/8$



Slide 5.7

V.A. discretas – Função de distribuição

Função de distribuição ($F(y)$)
Associa a cada valor de $y \in \Re$ a probabilidade de Y ser menor ou igual a y (também conhecida como probabilidade acumulada)

$$F(y) = \text{Probabilidade}(Y \leq y) = P(Y \leq y) = \sum_{y_k \leq y} p(y_k)$$

Propriedades:

(1) $F(y)$ é crescente

(2) $F(-\infty) = 0$

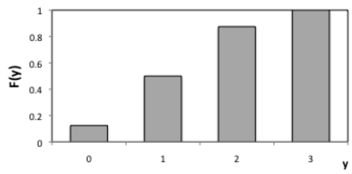
(3) $F(+\infty) = 1$

(4) $P(a < Y \leq b) = F(b) - F(a)$

Exemplo – Lançamento da moeda E-C três vezes ao ar

Y: número de E's em cada sequência

y	0	1	2	3
$p(y)$	$1/8$	$3/8$	$3/8$	$1/8$
$F(y)$	$1/8$	$4/8$	$7/8$	1



Slide 5.8

4.2.2 Exercícios

Um dado é lançado 3 vezes. Seja X uma variável aleatória que conta o número de vezes que sai uma face PAR.

a) Identifique a experiência aleatória e o espaço amostral associado a esta situação.
b) Determine a função de distribuição de probabilidade de X.
c) Determine a probabilidade de sair face PAR em pelo menos dois dos lançamentos.
d) Suponha que se sabe que num dos lançamento não saiu face PAR. Determine a probabilidade de sair face PAR em pelo menos dois dos lançamentos.

Mostre que a função de probabilidade $p(y)$ satisfaz as seguintes propriedades:

a) $p(y) \geq 0, \, y \in S$
b) $\sum_{y \in S} p(y) = 1$
c) $\sum_{y \in A} p(y) = P(Y \in A), \, A \subseteq S$

Slide 5.9

4.3 Distribuições Contínuas

- Nas v.a. contínuas, a probabilidade distribui-se de forma contínua por um intervalo real

- A probabilidade associada a um ponto é nula
(se assim não fosse, como temos um número infinito de pontos, considerando todos os pontos a probabilidade daria infinito)

- Mas é possível calcular a probabilidade associada a um intervalo finito:

$$\Delta P = P(x < X < x + \Delta x)$$

Slide 5.10

- À medida que vamos considerando Δx tão pequeno quanto queiramos, $\frac{\Delta P}{\Delta x}$ aproxima-se da densidade de probabilidade da variável X em torno de x, isto é:

$$\lim_{\Delta x \to 0} \frac{\Delta P}{\Delta x} = f(x)$$

- Usando o conceito de derivada:

$$\begin{aligned} f(x) &= \frac{\mathrm{d}P(x < X < x + \mathrm{d}x)}{\mathrm{d}x} \\ \mathrm{d}P(x < X < x + \mathrm{d}x) &= f(x)\mathrm{d}x \\ P(a \leq X \leq b) &= \int_a^b f(x)\mathrm{d}x \end{aligned}$$

Nota:
$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$, já que a probabilidade associada a um ponto é nula ($P(X = a) = P(X = b) = 0$)

Slide 5.11

4.3.1 Função densidade de probabilidade e função de distribuição

V.A. contínuas – Função densidade de probabilidade

Função densidade de probabilidade ($f(x)$)
Função que associa a cada valor particular x a concentração de probabilidade por unidade da variável X no ponto $X = x$

$$f(x): \quad P(a \leq X \leq b) = \int_a^b f(x)dx$$

Propriedades:

(1) $f(x) \geq 0, \ \forall_{x \in \Re}$
(2) $\int_{-\infty}^{+\infty} f(x) = 1$

Nota: $f(x)dx$ = probabilidade de $X \in [x, x + dx] \quad \neq \quad P(X = x) = 0$

Slide 5.12

Exemplo
Considere a v.a. T que representa o tempo sem avarias (ou entre avarias) expresso em dias de um determinado equipamento. A função densidade de probabilidade é dada pela seguinte expressão:

$$f(t) = \begin{cases} 0, & \text{se } t \leq 0 \\ 0.5\,e^{-0.5t}, & \text{se } t \geq 0 \end{cases}$$

- Calcule a probabilidade do equipamento funcionar sem avarias entre 1 e 3 dias:

$$P(1 < T < 3) = \int_1^3 f(t)\,dt = \int_1^3 0.5\,e^{-0.5t}\,dt = e^{-0.5} - e^{-1.5} = 0.3834$$

- Verifique que $f(t)$ satisfaz a condição $\int_{-\infty}^{+\infty} f(t)\,dt = 1$

$$\int_{-\infty}^{+\infty} f(t)\,dt = \int_{-\infty}^{0} 0\,dt + \int_0^{+\infty} 0.5\,e^{-0.5t}\,dt = 0 + e^0 - e^{-\infty} = 1 - 0 = 1$$

Slide 5.13

V.A. contínuas – Função de distribuição

Função de distribuição ($F(x)$)
Associa a cada valor de $x \in \Re$ (conjunto dos números reais) a probabilidade de X ser menor ou igual a x

Nota: também conhecida como probabilidade acumulada

$$F(x) \;=\; P(X \leq x) \;=\; \int_{-\infty}^{x} f(u)du$$

Propriedades:

(1) $F(x)$ é monótona crescente ($x_2 > x_1 \Rightarrow F(x_2) \geq F(x_1)$)

(2) $F(-\infty) = 0$ e $F(+\infty) = 1$

(3) $P(a \leq X \leq b) = F(b) - F(a)$

(4) $F(x) \geq 0 \quad \forall_x$

(5) $\frac{dF(x)}{dx} = f(x)$

Slide 5.14

Exemplo
Considere de novo a v.a. T do exemplo do slide 4-11. Obtenha a função de distribuição ($F(t)$):

- Para $t \leq 0$:
$$F(t) = P(T \leq t) = \int_{-\infty}^{t} 0\,dt = 0$$
- Para $t > 0$:
$$F(t) = P(T \leq 0) = F(0) + \int_{0}^{t} f(t)\,dt = \int_{0}^{t} 0.5\,e^{-0.5t}\,dt = 1 - e^{-0.5t}$$
- Ou seja, a função de distribuição de probabilidade ($F(t)$) é dada por:
$$F(t) = \begin{cases} 0, & \text{se } t \leq 0 \\ 1 - e^{-0.5t}, & \text{se } t \geq 0 \end{cases}$$

Slide 5.15

4.3.2 Exercícios

Seja $f(x) = -\ln(x),\ 0 < x < 1$

1. Esboce o gráfico de $f(x)$
2. Mostre que $f(x)$ é uma função densidade de probabilidade
3. Calcule $P(1/3 \leq X \leq 1/2)$

Mostre que a função de distribuição de probabilidade $F(x)$ satisfaz as seguintes propriedades:

1. $F(x)$ é monótona crescente
2. $F(-\infty) = 0$ e $F(+\infty) = 1$
3. $P(a \leq X \leq b) = F(b) - F(a)$
4. $F(x) \geq 0$, $\forall x \in \Re$ é monótona crescente
5. $\frac{dF(x)}{dx} = f(x)$

Slide 5.16

4.4 Representação de Populações

- *Conhecida* a priori *a população completa*, então pode-se fazer corresponder às frequências relativas valores de probabilidade e obter assim uma *função de probabilidade* (assumindo o caso discreto)

- Quando não é possível caracterizar com rigor absoluto a população completa (caso mais habitual) *a função de probabilidade assume o papel de modelo aproximado* para a distribuição dos elementos da população

- Procedimento para a constituição deste modelo: obter e caracterizar completamente uma amostra significativa da população e usar as *frequências relativas obtidas como um modelo aproximado da população*

- Note-se que para uma determinada população, enquanto a distribuição de probabilidade é fixa, já *as distribuições de frequências relativas variam de amostra para amostra*

Slide 5.17

4.5 Parâmetros das Distribuições

Parâmetros das Distribuições (ou Parâmetros Populacionais)

Permitem caracterizar distribuições (ou populações)

Parâmetros de localização: média (μ), mediana (η), moda (ζ)

Parâmetros de dispersão: desvio absoluto médio (δ), variância (σ^2), desvio-padrão (σ)

Outros parâmetros: momentos ordinários na origem (μ_i'), momentos centrados (μ_i), coeficiente de *Skewness* (assimetria) (γ_1), coeficiente de *Kurtosis* (achatamento) (γ_2)

Notas

- Os parâmetros desempenham em relação às distribuições populacionais um papel idêntico ao que as estatísticas desempenhavam em relação às distribuições amostrais
- Os parâmetros de uma população são fixos em contraste com as estatísticas que variam de amostra para amostra
- Os parâmetros populacionais representam-se por letras do alfabeto grego (estatísticas representam-se por letras do alfabeto latino)

Slide 5.18

4.5.1 Parâmetros de localização

Média de uma V.A.

Média (ou valor esperado ou esperança matemática) (μ)

$$\mu_Y = E(Y) = \sum_{i=1}^{n} y_i \cdot p(y_i) \qquad \text{(variáveis discretas)}$$

$$\mu_X = E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\, dx \qquad \text{(variáveis contínuas)}$$

Propriedades do valor esperado:

1. Se $X = c$ (constante) então $E(X) = c$
2. $E(c \cdot X) = c \cdot E(X)$

Slide 5.19

Outros parâmetros de localização

Mediana (η)

Distribuições discretas: Primeiro valor para o qual a função de distribuição é maior ou igual a 0.5 ($F(x) \geq 0.5$)

Distribuições contínuas: Valor de x para o qual a função de distribuição é igual a 0.5

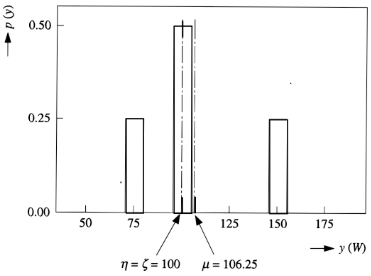
- $\eta_X = x$, se $F(x) = 0.5$ só tem uma solução
- $\eta_X = (x_{min} + x_{max})/2$, no caso contrário (onde x_{min} e x_{max} são o mínimo e máximo do conjunto de soluções)

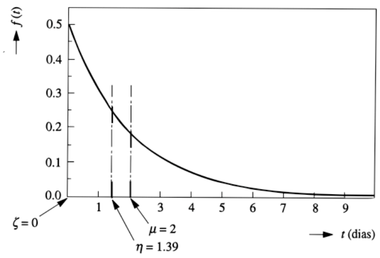
Moda (ζ)
Valor da variável para o qual é máxima a:

- função de probabilidade (*distribuições discretas*)
- função densidade de probabilidade (*distribuições contínuas*)

Slide 5.20

Média, moda e mediana de uma V.A. – Exemplo





E se tivéssemos gráficos com funções de distribuição ($F(x)$) ...

Slide 5.21

4.5.2 Parâmetros de dispersão

Desvio absoluto médio (δ)
$\delta_X = E\{|X - \mu_X|\}$

Variância (σ^2)
$\sigma_X^2 = Var(X) = E\{(X - \mu_X)^2\}$

Vindo,

$$\sigma_Y^2 = Var(Y) = \sum_{i=1}^{n} (y_i - \mu_Y)^2 \cdot p(y_i) \qquad \text{(variáveis discretas)}$$

$$\sigma_X^2 = Var(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f(x)\, dx \qquad \text{(variáveis contínuas)}$$

Desvio-padrão (σ)
$\sigma_X = +\sqrt{\sigma_X^2} = +\sqrt{Var(X)}$

Slide 5.22

Propriedades da variância:

1. $\sigma^2_{(X+c)} = Var(X+c) = Var(X)$ com $c =$ constante

$$\left| Var(X+c) = E\{(X + c - (\mu_X + c))^2\} = E\{(X - \mu_X)^2\} = Var(X) \right.$$

2. $\sigma^2_{(c\cdot X)} = Var(c \cdot X) = c^2 \cdot Var(X)$ com $c =$ constante

$$\left| \begin{aligned} Var(c \cdot X) &= E\{(c \cdot X - c \cdot \mu_X)^2\} = E\{c^2 \cdot (X - \mu_X)^2\} = \\ &= c^2 \cdot E\{(X - \mu_X)^2\} = c^2 \cdot Var(X) \end{aligned} \right.$$

3. $\sigma_X^2 = Var(X) = E(X^2) - E^2(X)$ (fórmula de cálculo simplificada)

$$\left| \begin{aligned} Var(X) &= E\{(X - \mu_X)^2\} = E(X^2 - 2X\mu_X + \mu_X^2) = E(X^2) - 2\mu_X E(X) + \\ &+ \mu_X^2 = E(X^2) - 2E^2(X) + E^2(X) = E(X^2) - E^2(X) \end{aligned} \right.$$

Nota: $Var(2X) \neq Var(X) + Var(X)$
(dado que $Var(2X) = 4Var(X)$ e que $Var(X) + Var(X) = 2Var(X)$)

Slide 5.23

4.5.3 Outros parâmetros – Momentos

Momentos ordinários de ordem k (μ_k')

$$\mu_k' = \sum_{x_i} y_i^k \cdot p(y_i) \qquad \text{(v.a. discretas)}$$

$$\mu_k' = \int_{-\infty}^{+\infty} x^k \cdot f(x)\, dx \qquad \text{(v.a. contínuas)}$$

Momentos centrados de ordem k (μ_k)

$$\mu_k = \sum_{y_i} [y_i - E(Y)]^k \cdot p(y_i) \qquad \text{(v.a. discretas)}$$

$$\mu_k = \int_{-\infty}^{+\infty} [x - E(X)]^k \cdot f(x)\, dx \qquad \text{(v.a. contínuas)}$$

Notas:

- Valor esperado e variância são casos particulares de momentos de uma v.a.
- Os momentos centrados e ordinários de uma v.a. estão relacionados (por exemplo: $\mu_2 = \mu_2' - (\mu_1')^2$) Slide 5.24

Significado de alguns momentos:

$\mu_1 = 0$
μ_1' – Valor esperado (μ)
μ_2 – Variância (σ^2)
μ_3 – Assimetria
μ_4 – Kurtose (achatamento)

Coeficiente populacional de assimetria (γ_1)

$$\gamma_1 = \frac{\mu_3}{\sigma^3}$$

Coeficiente populacional de kurtose (γ_2)

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$$

Slide 5.25

4.5.4 Exemplo – Quiosque

A procura de uma revista importada num quiosque (Y) segue a seguinte distribuição:

y	0	1	2	3	4	5	6	≥ 7
$p_Y(y)$	0.0467	0.1866	0.311	0.2765	0.1382	0.0369	0.0041	0

1. Calcule o valor esperado e o desvio padrão da procura dessa revista.
2. Calcule os coeficientes de populacionais de assimetria (γ_1) e de kurtose (γ_2) da procura dessa revista. Interprete os resultados obtidos.
3. Sabendo que foi enviada uma remessa de 5 exemplares da revista para o quiosque, calcule o valor esperado e o desvio padrão do número de:
   a. exemplares vendidos (Z).
   b. exemplares que não foram vendidos por falta de procura (W).
   c. clientes que tentaram comprar a revista e não o conseguiram por excesso de procura (V).

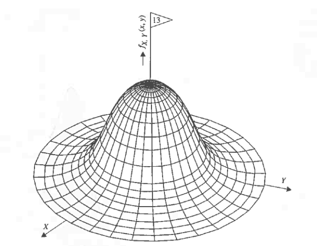
Soluções: 1. $\mu_Y = 2.40$, $\sigma_Y = 1.20$, 2. $\gamma_{1Y} = 0.1667$, $\gamma_{2Y} = -0.3054$, 3.a) $\mu_Z = 2.3959$, $\sigma_Z = 1.19$, 3.b) $\mu_W = 2.60$, $\sigma_W = 1.19$, 3.c) $\mu_V = 0.0041$, $\sigma_V = 0.064$ (nota: $\mu_Y = \mu_Z + \mu_V$, mas $\sigma_Y \neq \sigma_Z + \sigma_V$, porquê?) Slide 5.26

4.6 Distribuições conjuntas de probabilidade

Surgem em experiências aleatórias que envolvem simultaneamente várias variáveis

$\Rightarrow$ estamos interessados nas propriedades individuais de cada variável
$\Rightarrow$ mas também nas relações entre elas (propriedades *conjuntas*)

Elementos da população	Altura [m]	Peso [kg]
A	1.73	75
B	1.76	75
C	1.76	75
D	1.78	76
E	1.78	77
F	1.78	79
G	1.82	81



- Variáveis aleatórias bidimensionais (discretas ou contínuas)
- Conceito facilmente generalizável a variáveis n-dimensionais
- Adaptável a situações que misturem V.A. discretas e contínuas (V.A. mistas)

Slide 5.27

4.6.1 Distribuições conjuntas de probabilidade

Distribuições conjuntas – Caso discreto

Função de probabilidade conjunta ($p_{XY}(x,y)$)
Função que associa a cada par de valores particulares de X e Y a probabilidade de ocorrer *simultaneamente* $X = x$ *e* $Y = y$

$$p_{XY}(x,y) = \text{Probabilidade}(X = x \wedge Y = y) = P(X = x, Y = y),\ \forall_{(x,y)}$$

Representação tabular de $p_{XY}(x,y)$:

	X [kg]	**75**	**76**	**77**	**79**	**81**
	1.73	1/7	0	0	0	0
Y	**1.76**	2/7	0	0	0	0
[m]	**1.78**	0	1/7	1/7	1/7	0
	1.82	0	0	0	0	1/7

Propriedades:

- $p_{XY}(x,y) \geq 0,\ \forall_{(x,y)}$
- $\sum_x \sum_y p_{XY}(x,y) = 1$

Função de distribuição conjunta ($F_{XY}(x,y)$)

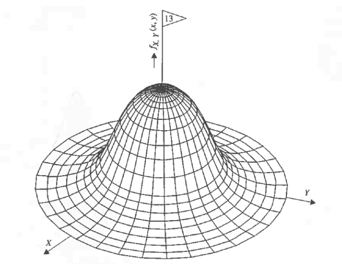
$$F_{XY}(x,y) = P(X \leq x, Y \leq y) = \sum_{x' \leq x} \sum_{y' \leq y} p_{XY}(x',y')$$

Slide 5.28

Distribuições conjuntas – Caso contínuo

Função de densidade de probabilidade conjunta ($f_{XY}(x,y)$)

$$\text{Probabilidade}[(X,Y) \in D] = \int\int_D f_{XY}(x,y)\, dx\, dy$$



Propriedades:

- $f_{XY}(x,y) \geq 0,\ \forall_{i,j}$
- $\int\int_{\Re^2} f_{XY}(x,y)\, dy\, dx = 1$

Função de distribuição conjunta ($F_{XY}(x,y)$)

$$F_{XY}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x,y)\, dx\, dy$$

Slide 5.29

4.6.2 Distribuições marginais de probabilidade

Funções de probabilidade marginais (caso discreto)

$$p_X(x) = \sum_y p_{XY}(x,y) \qquad p_Y(y) = \sum_x p_{XY}(x,y)$$

Funções densidade de probabilidade marginais (caso contínuo)

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy \qquad f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dx$$

Funções de distribuição marginais (iguais no caso discreto e contínuo)

$$F_X(x) = F_{XY}(x,+\infty) \qquad F_Y(y) = F_{XY}(+\infty,y)$$

Notas:

- As distribuições marginais obtêm-se "somando" (ou integrando) sobre a outra variável
- Não se pode obter a distribuição conjunta a partir das distribuições marginais (impossível recuperar as propriedades conjuntas das V.A.)

Slide 5.30

4.6.3 Distribuições condicionais de probabilidade

Variáveis discretas:

$$p_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)} = \frac{p_{XY}(x,y)}{\sum_x p_{XY}(x,y)}$$
$$p_{Y|X}(y|x) = \frac{p_{XY}(x,y)}{p_X(x)} = \frac{p_{XY}(x,y)}{\sum_y p_{XY}(x,y)}$$

Variáveis contínuas:

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

Notas:

- As distribuições condicionais estão definidas apenas para os valores em que as funções (densidade) de probabilidade marginais são positivas $\left(p_X(x_i) > 0,\ p_Y(y_j) > 0,\ f_X(x) > 0,\ f_Y(y) > 0\right)$
- As distribuições condicionais são funções que dependem *simultaneamente* de X e Y

Slide 5.31

Valor esperado condicionado

Variáveis discretas:

$$E(X|Y=y) = \sum_x x \cdot p_{X|Y}(x|y) \qquad E(Y|X=x) = \sum_y y \cdot p_{Y|X}(y|x)$$

Variáveis contínuas:

$$E(X|Y=y) = \int_{-\infty}^{+\infty} x \cdot f_{X|Y}(x|y)\,dx \qquad E(Y|X=x) = \int_{-\infty}^{+\infty} y \cdot f_{Y|X}(y|x)\,dy$$

Nota: O valor esperado condicionado de $E(X|Y=y)$ não é uma constante, mas uma função que depende do valor de y

Propriedades:

1. $E[E(X|Y=y)] = E(X) \quad \wedge \quad E[E(Y|X=x)] = E(Y)$
2. Se X e Y forem variáveis independentes, então $E(X|Y=y) = E(X) \quad \wedge \quad E(Y|X=x) = E(Y)$

Slide 5.32

4.6.4 Exemplo – Livraria

A livraria do Sr. Manuel vende regularmente livros estrangeiros em língua inglesa e língua francesa. O número de vezes que um cliente compra livros em inglês (X) e em francês (Y) semanalmente estão indicados na tabela seguinte (por exemplo: há 10 clientes que compram 3 livros em inglês e 2 livros em francês):

		X				
		0	**1**	**2**	**3**	**4**
	0	2	3	3	2	0
Y	**1**	4	12	12	6	2
	2	3	9	12	10	6
	3	1	6	3	2	2

Determine:

1. a função de probabilidade conjunta ($p_{XY}(x,y)$)
2. a função de distribuição conjunta ($F_{XY}(x,y)$)

Slide 5.33

1) Função de probabilidade conjunta

		X				
$p_{XY}(x,y)$		**0**	**1**	**2**	**3**	**4**
	0	0.02	0.03	0.03	0.02	0.00
Y	**1**	0.04	0.12	0.12	0.06	0.02
	2	0.03	0.09	0.12	0.10	0.06
	3	0.01	0.06	0.03	0.02	0.02

2) Função de distribuição conjunta

$F_{XY}(x,y)$		X: **0**	**1**	**2**	**3**	**4**
	0	0.02	0.05	0.08	0.10	0.10
Y	**1**	0.06	0.21	0.36	0.44	0.46
	2	0.09	0.33	0.60	0.78	0.86
	3	0.10	0.40	0.70	0.90	1.00

Slide 5.34

3. Determine as funções de probabilidade marginais ($p_X(x)$ e $p_y(y)$)

$p_{XY}(x,y)$		X: **0**	**1**	**2**	**3**	**4**	$p_Y(y)$
	0	0.02	0.03	0.03	0.02	0.00	0.10
Y	**1**	0.04	0.12	0.12	0.06	0.02	0.36
	2	0.03	0.09	0.12	0.10	0.06	0.40
	3	0.01	0.06	0.03	0.02	0.02	0.14
$p_X(x)$		0.10	0.30	0.30	0.20	0.10	1

4. Determine as funções de distribuição marginais ($F_X(x)$ e $F_Y(y)$)

$F_{XY}(x,y)$		X: **0**	**1**	**2**	**3**	**4**	$F_Y(y)$
	0	0.02	0.05	0.08	0.10	0.10	0.10
Y	**1**	0.06	0.21	0.36	0.44	0.46	0.46
	2	0.09	0.33	0.60	0.78	0.86	0.86
	3	0.10	0.40	0.70	0.90	1.00	1.00
$F_X(x)$		0.10	0.40	0.70	0.90	1.00	1.00

Slide 5.35

5. Determine a função condicional de probabilidade $p_{Y|X}(y|x)$
6. Determine o valor esperado de Y condicionado a X, $E(Y|X=x)$

$p_{XY}(x,y)$		X: **0**	**1**	**2**	**3**	**4**	$p_Y(y)$
	0	0.02	0.03	0.03	0.02	0.00	0.10
Y	**1**	0.04	0.12	0.12	0.06	0.02	0.36
	2	0.03	0.09	0.12	0.10	0.06	0.40
	3	0.01	0.06	0.03	0.02	0.02	0.14
$p_X(x)$		0.10	0.30	0.30	0.20	0.10	1
	0	0.20	0.10	0.10	0.10	0.00	
Y	**1**	0.40	0.40	0.40	0.30	0.20	$p_{Y|X}(y|x)$
	2	0.30	0.30	0.40	0.50	0.60	
	3	0.10	0.20	0.10	0.10	0.20	
$E(Y|X=x)$		1.3	1.6	1.5	1.6	2.0	

$$E(Y|X=0) = \sum_y y \cdot p_{Y|X=0}(0,y) = 0 \times 0.20 + 1 \times 0.40 + 2 \times 0.30 + 3 \times 0.10 = 1.3$$

$$E(Y|X=1) = \sum_y y \cdot p_{Y|X=1}(1,y) = 0 \times 0.10 + 1 \times 0.40 + 2 \times 0.30 + 3 \times 0.30 = 1.6$$

$$E(Y|X=2) = \sum_y y \cdot p_{Y|X=2}(2,y) = 0 \times 0.10 + 1 \times 0.40 + 2 \times 0.40 + 3 \times 0.10 = 1.5$$

$$E(Y|X=3) = \sum_y y \cdot p_{Y|X=3}(3,y) = 0 \times 0.10 + 1 \times 0.30 + 2 \times 0.50 + 3 \times 0.10 = 1.6$$

$$E(Y|X=4) = \sum_y y \cdot p_{Y|X=4}(4,y) = 0 \times 0.00 + 1 \times 0.20 + 2 \times 0.60 + 3 \times 0.20 = 1.8$$

Slide 5.36

4.7 Independência entre variáveis aleatórias

Variáveis aleatórias independentes

- Duas V.A. (X,Y) dizem-se independentes se, para quaisquer valores (x,y) que elas possam tomar, os acontecimentos $X = x$ e $Y = y$ também forem independentes

$$\begin{cases} \text{Probabilidade}(X = x|Y = y) = \text{Probabilidade}P(X = x) \\ \text{Probabilidade}(Y = y|X = x) = \text{Probabilidade}(Y = y) \end{cases} ,\forall_{(x,y)}$$

- Ou seja, (X,Y) são duas variáveis aleatórias independentes se:

$$p_{X|y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)} = p_X(x),\ \forall_{(x,y)}$$

ou

$$p_{Y|x}(y|x) = \frac{p_{XY}(x,y)}{p_X(x)} = p_Y(y),\ \forall_{(x,y)}$$

ou

$$p_{XY}(x,y) = p_X(x).p_Y(y),\ \forall_{(x,y)}$$

Slide 5.37

- Para *variáveis contínuas*:

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

ou

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

ou

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

Condições necessárias e suficientes de independência entre n V.A.

- *variáveis discretas*: $p_{Y_1Y_2\ldots Y_N}(y_1,y_2,\ldots,y_N) = p_{Y_1}(y_1)\cdot p_{Y_2}(y_2)\cdot\ldots\cdot p_{Y_N}(y_N), \quad \forall_{(y_1,y_2,\ldots,y_N)}$
- *variáveis contínuas*: $f_{X_1X_2\ldots X_N}(x_1,x_2,\ldots,x_N) = f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot\ldots\cdot f_{X_N}(x_N), \quad \forall_{(x_1,x_2,\ldots,x_N)}$

Slide 5.38

4.7.1 Exemplo – Livraria (cont.)

7. Verifique se X e Y são variáveis aleatórias independentes.

$$p_{XY}(x,y) \stackrel{?}{=} p_X(x).p_Y(y)\ ,\ \forall_{(x,y)}$$
$$p_{XY}(0,0) = \mathbf{0.02} \quad \neq \quad p_X(0).p_Y(0) = 0.10\times 0.10 = \mathbf{0.01}$$

$\Longrightarrow$ Logo *não são independentes*

- para provar que não são independentes basta encontrar um caso em que a igualdade não se verifique
- para provar que são independentes é necessário verificar todos os casos

  Alternativamente, pode-se verificar se:

$$p_{Y|X}(y|x) \stackrel{?}{=} p_Y(y)$$

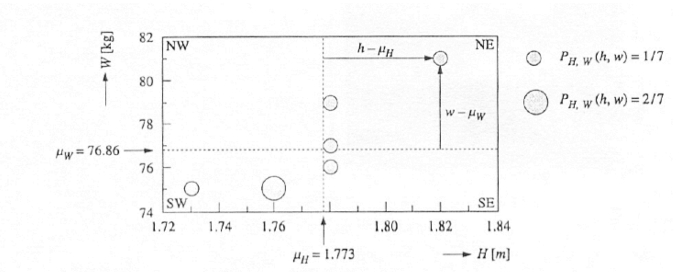
  Observando a tabela com as respostas às questões 5. e 6. do exemplo facilmente se constata que $p_{Y|X}(y|x) \neq p_Y(y)$.

$\Longrightarrow$ Logo as variáveis X e Y não são independentes

Slide 5.39

4.8 Covariância e correlação

- Variáveis aleatórias *não* independentes podem estar relacionadas de diferentes formas
- Relacionamento *linear* é a forma mais simples de relacionamento (pode ser utilizada como aproximação das restantes)
- *Covariância* e *correlação* são medidas do grau de relacionamento linear entre duas variáveis



Slide 5.40

Covariância populacional (γ_{XY})

- Média pesada dos produtos dos desvios das variáveis face às respectivas médias

$$\gamma_{XY} = Cov(X,Y) = E\{[X-\mu_X][Y-\mu_Y]\}$$

Caso discreto: $\gamma_{XY} = \sum_x \sum_y [x-\mu_X]\cdot[y-\mu_Y]\cdot p_{XY}(x,y)$

Caso contínuo: $\gamma_{XY} = \int\int_{\Re^2} [x-\mu_X]\cdot[y-\mu_Y]\cdot f_{XY}(x,y)\,dy\,dx$

- A covariância tem o inconveniente de ser afectada pelas dimensões das variáveis

$\Rightarrow$ *Coeficiente de correlação populacional* (medida adimensional)

Slide 5.41

Coeficiente de correlação populacional (ρ_{XY})

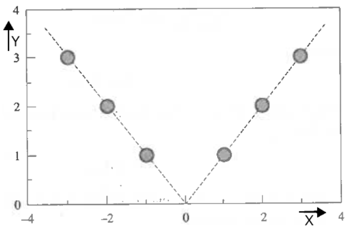
- Medida *padronizada* do grau de relacionamento linear entre duas variáveis

$$\rho_{XY} = \frac{\gamma_{XY}}{\sigma_X\cdot\sigma_Y}$$

- $\rho_{XY} = \pm 1 \;\Rightarrow\;$ Relacionamento linear perfeito entre X e Y
- $\rho_{XY} = 0 \;\Rightarrow\;$ Não há relacionamento linear entre X e Y

Notas:

- $-1 \le \rho_{XY} \le 1$
- (X,Y) são V.A. independentes $\Rightarrow \rho_{XY} = 0$
  $\rho_{XY} = 0 \not\Rightarrow (X,Y)$ sejam independentes
  (Exemplo: $Y = |X| \;\rightarrow\; \gamma_{XY} = \rho_{XY} = 0$)



Slide 5.42

4.8.1 Exemplo – Livraria (cont.)

8. Determine a covariância entre X e Y (γ_{XY})

$$\mu_X = E(X) = \sum_x x\cdot p_X(x) = 0\times 0.10 + 1\times 0.30 + 2\times 0.30 + 3\times 0.20 + 4\times 0.10 = 1.9$$

$$\mu_Y = E(Y) = \sum_y y\cdot p_Y(y) = 0\times 0.10 + 1\times 0.36 + 2\times 0.40 + 3\times 0.14 = 1.58$$

$$\gamma_{XY} = \sum_x\sum_y [x-\mu_X]\cdot[y-\mu_Y]\cdot p_{XY}(x,y) = (0-1.9)\times(0-1.58)\times 0.02 +$$

$$+(0-1.9)\times(1-1.58)\times 0.04+\ldots+(4-1.9)\times(3-1.58)\times 0.02=0.138$$

Alternativamente, γ_{XY} pode ser calculado recorrendo à seguinte fórmula:

$$\gamma_{XY}=\mu_{XY}-\mu_X\cdot\mu_Y=E(XY)-E(X)\cdot E(Y)$$

$$\mu_{XY}=E(XY)=\sum_x\sum_y(x\cdot y)\cdot p_{XY}(x,y)=(0\times 0)\times 0.02+\ldots+$$

$$+(1\cdot 1)\times 0.12+(1\cdot 2)\times 0.12+\ldots+(3\cdot 4)\times 0.02=3.14$$

Vindo,

$$\gamma_{XY}=E(XY)-E(X)\cdot E(Y)=3.14-1.9\times 1.58=0.138$$

Slide 5.43

9. Determine o coeficiente de correlação entre X e Y (ρ_{XY})

$$\mu_{X^2}=E(X^2)=\sum_i x^2\cdot p_X(x_i)=0^2\times 0.10+1^2\times 0.30+2^2\times 0.30+$$

$$+3^2\times 0.20+4^2\times 0.10=4.9$$

$$\mu_{Y^2}=E(Y^2)=\sum_y y^2\cdot p_Y(y)=0^2\times 0.10+1^2\times 0.36+2^2\times 0.40+$$

$$+3^2\times 0.14=3.22$$

$$\sigma_X^2=E(X^2)-E^2[X]=4.9-1.9^2=1.29$$
$$\sigma_Y^2=E(Y^2)-E^2[Y]=3.22-1.58^2=0.7236$$

$$\rho_{XY}=\frac{\gamma_{XY}}{\sigma_X\sigma_Y}=\frac{0.138}{\sqrt{1.29}\times\sqrt{0.7236}}=0.1428$$

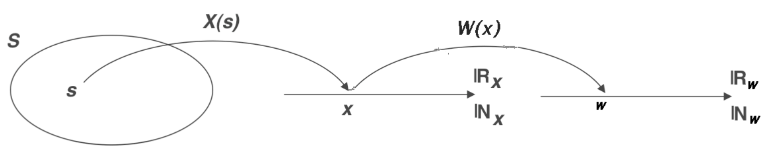
Slide 5.44

4.9 Variáveis Transformadas

Variável transformada
Aplicação do domínio da v.a. X na v.a. W:

$$W=\Phi(X)$$

Probabilidade do acontecimento s no espaço amostral $S =$ = Probabilidade do acontecimento equivalente no domínio da v.a. $X =$ = Probabilidade do acontecimento equivalente no domínio da v.a. W



Slide 5.45

4.9.1 Transformação de variáveis aleatórias

Transformação de V.A. discretas
Conhecidas:

- a função de probabilidade da v.a. Y ($p_Y(y_i)$) e
- a função de transformação ($W=\Phi(Y)$)

A função de probabilidade da v.a. W é dada por:

$$p_W(w_j)=P(W=w_j)=\sum_i P(Y=y_i),\quad \text{com } y_i:\Phi(y_i)=w_j$$

Exemplo

- Y: Potência de uma lâmpada
- $p_Y(y)$: $\quad p_Y(75) = 0.25, \quad p_Y(100) = 0.50, \quad p_Y(150) = 0.25$
- $W = \Phi(Y) = (Y - 100)/25 = (Y/25) - 4$

$\Rightarrow$ Função de probabilidade da v.a. W ($p_W(w_j)$):

$$\begin{array}{lcl} y_1 = 75 : w_1 = (75/25) - 4 = -1 & & p_W(-1) = p_Y(75) = 0.25 \\ y_2 = 100 : w_2 = (100/25) - 4 = 0 & \rightarrow & p_W(0) = p_Y(100) = 0.50 \\ y_3 = 150 : w_3 = (150/25) - 4 = 2 & & p_W(2) = p_Y(150) = 0.25 \end{array}$$

Slide 5.46

Transformação de V.A. contínuas

Nota: tópico não avaliado, apresentado apenas a título ilustrativo

Conhecidas:

- a função densidade de probabilidade da v.a. X ($f_X(x)$)
- e a função de transformação ($W = \Phi(X)$)

A função de probabilidade da v.a. W pode ser obtida por

- Se $\Phi(x)$ é monótona pode-se obter a função inversa $\Phi^{-1}(w)$

$$f_W(w) = \left. \frac{f_X(x)}{\left|\frac{dw}{dx}\right|} \right|_{x=\Phi^{-1}(w)}$$

- Senão é necessário dividir $\Phi(x)$ em regiões monótonas (K):

$$f_W(w) = \sum_i \left. \frac{f_X(x)}{\left|\frac{dw}{dx}\right|} \right|_{\substack{x=\Phi^{-1}(w) \\ x \in \text{ região } i \\ w \in \text{ região } i}}, \ (i = 1, \ldots, K)$$

Slide 5.47

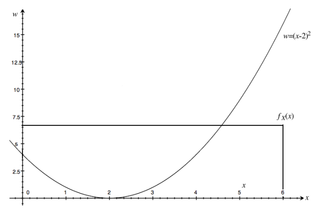
Transformação de V.A. contínuas – Exemplo

Nota: tópico não avaliado, apresentado apenas a título ilustrativo

X: atraso de um comboio (min)
(atraso máximo: 6 min)

$$f_X(x) = \begin{cases} 1/6, & 0 \le X \le 6 \\ 0, & \text{outros valores} \end{cases}$$

$W = \Phi(X) = (X-2)^2$



Como $\Phi(x)$ não é monótona é necessário dividir em duas regiões: $x < 2$ e $x \ge 2$, vindo:

$$f_W(w) = \sum_{i=1}^{2} \left. \frac{f_X(x)}{\left|\frac{dw}{dx}\right|} \right|_{\substack{x=\Phi^{-1}(w) \\ x \in \text{ região } i \\ w \in \text{ região } i}} = \left. \frac{1/6}{|2 \cdot (x-2)|} \right|_{\substack{x=2-\sqrt{w} \\ 0<x<2 \\ 0<w<4}} + \left. \frac{1/6}{|2 \cdot (x-2)|} \right|_{\substack{x=2+\sqrt{w} \\ 2<x<6 \\ 0<w<16}}$$

$$f_W(w) = \begin{cases} \frac{1/3}{|2-\sqrt{w}-2|} + \frac{1/3}{|2+\sqrt{w}-2|} = \frac{1/3}{\sqrt{w}} + \frac{1/3}{\sqrt{w}} = \frac{2}{3 \cdot \sqrt{w}}, \text{ para } 0 \le w \le 4 \\ \frac{1/3}{|2+\sqrt{w}-2|} = \frac{1}{3 \cdot \sqrt{w}}, \text{ para } 4 < w \le 16 \end{cases}$$

Slide 5.48

4.9.2 Momentos populacionais definidos como valores esperados

Momentos definidos como valores esperados de variáveis transformadas

Seja Z uma v.a. inteira e definindo as transformações:

$$W_1 = Z^i \quad \text{e} \quad W_2 = (Z - \mu_Z)^i$$

Os valores esperados de W_1 e W_2 são dados por:

$$E(W_1) = \sum_{w_1} w_1 \cdot p_{W_1}(w_1) = \sum_z z^i \cdot p_Z(z) = E(Z^i)$$

$$E(W_2) = \sum_{w_2} w_2 \cdot p_{W_2}(w_2) = \sum_{z} (z-\mu_Z)^i \cdot p_Z(z) = E[(Z-\mu_Z)^i]$$

Nota: facilmente generalizável ao caso de variáveis aleatórias contínuas

Momentos de ordem i da v.a. Z

$$\mu_i'(Z) = E(Z^i) \qquad\qquad \mu_i = E[(Z-\mu_Z)^i]$$

Slide 5.49

4.9.3 Cálculo do valor esperado e da variância de variáveis transformadas

- Começando por calcular a função (densidade) de probabilidade da variável transformada e calculando μ e σ^2 pelas suas definições (ver slide 4-47)

- Ou calculando, de forma exacta no caso de transformações lineares ou aproximada no caso de transformações não lineares, directamente a partir dos parâmetros da distribuição da variável original:

Transformações lineares (cálculo exacto):

Seja

$$W = \Phi(Z) = a + b \cdot Z$$

Então:

$$\mu_W = E(W) = E(a+b\cdot Z) = a + b\cdot E(Z) = a + b\cdot \mu_Z$$

$$\sigma_W^2 = Var(W) = Var(a+b\cdot Z) = b^2 \cdot Var(Z) = b^2 \cdot \sigma_Z^2$$

Slide 5.50

Nota: tópico não avaliado, apresentado apenas a título ilustrativo

Transformações não lineares (cálculo aproximado):

- μ_W e σ_W^2 dependem não só de μ_Z e σ_Z^2 mas também de momentos de ordem superior de Z
- **Simplificação:** aproximar as transformações não-lineares por uma relação linear, desenvolvendo $\Phi(Z)$ em série de Taylor em torno de μ_Z e tomando apenas os dois primeiros termos

$$\Phi(Z) \approx \Phi(\mu_Z) + (Z-\mu_Z)\left(\frac{\mathrm{d}\Phi}{\mathrm{d}Z}\right)_{Z=\mu_Z}$$

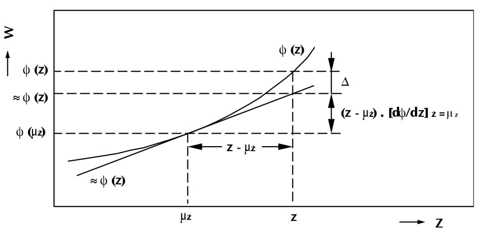
- μ_Z, $\Phi(\mu_Z)$ e $\left(\frac{\mathrm{d}\Phi}{\mathrm{d}Z}\right)_{Z=\mu_Z}$ são constantes $\longrightarrow$ relação linear, com:

$$a = \Phi(\mu_Z) - \mu_Z \cdot \left(\frac{\mathrm{d}\Phi}{\mathrm{d}Z}\right)_{Z=\mu_Z} \qquad b = \left(\frac{\mathrm{d}\Phi}{\mathrm{d}Z}\right)_{Z=\mu_Z}$$

- Substituindo no cálculo de μ_W e σ_W^2 quando $W = a + b\cdot Z$, vem:

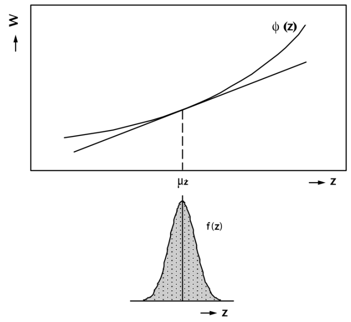
$$E(W) \approx \Phi(\mu_Z) \qquad Var(W) \approx \left(\frac{\mathrm{d}\Phi}{\mathrm{d}Z}\right)^2_{Z=\mu_Z} \cdot \sigma_Z^2$$

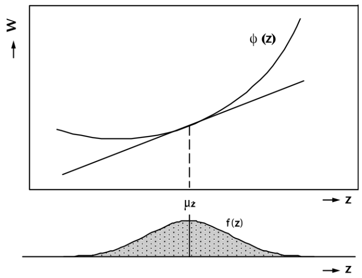
Slide 5.51



Qualidade da aproximação linear depende:

- da curvatura de $\Phi(Z)$ em torno de μ_Z (quanto menor melhor)
- e da dispersão da distribuição de Z (quanto menor melhor)





Slide 5.52

4.9.4 Exemplo – Quiosque (cont.)

A procura de uma revista importada num quiosque (Y) segue a seguinte distribuição:

y	0	1	2	3	4	5	6	≥ 7
$p_Y(y)$	0.0467	0.1866	0.311	0.2765	0.1382	0.0369	0.0041	0

4. Sabendo que:
   - o preço de venda de cada exemplar da revista é de 10 €;
   - o quiosque paga 8 € por cada exemplar recebido;
   - cada exemplar não vendido pode ser devolvido por 7 €;
   - de forma a minimizar a insatisfação dos clientes os donos do quiosque consideram uma penalização de 5 € por cada cliente que queira comprar um exemplar da revista e não o consiga.

   Determine o número de exemplares da remessa que maximiza o valor esperado do lucro (L) a obter com a venda de revistas.

Nota resolva esta questão de duas formas: 1. obtendo a função de probabilidade do lucro $p_L()$; 2. definindo o lucro como uma função do número de revistas vendidas, devolvidas e em falta.

Slide 5.53

A procura de uma revista importada num quiosque (Y) segue a seguinte distribuição:

y	0	1	2	3	4	5	6	≥ 7
$p_Y(y)$	0.0467	0.1866	0.311	0.2765	0.1382	0.0369	0.0041	0

5. Calcule o desvio padrão do lucro para o número de exemplares da remessa obtido na alínea anterior.
6. Concorda com a regra anterior de determinação do tamanho da remessa?

Soluções:
4. $\mu_{L0} = -12€$, $\mu_{L1} = -5.3736€$, $\mu_{L2} = -0.24€$, $\mu_{L3} = 2.4056€$, $\mu_{L4} = 2.8392€$, $\mu_{L5} = 2.1672€$, $\mu_{L6} = 1.2€ \longrightarrow$ remessa = 4 exemplares.
5. $\sigma_{L4} = 3.197€$.
6. Questão a ser analisada no trabalho #2 ...

Slide 5.54

4.9.5 Transformação de pares de variáveis aleatórias

- Dada a V.A. (X,Y) e uma função de transformação $\Phi(X,Y)$, fica definida uma nova variável aleatória

$$Z = \Phi(X,Y)$$

- A função (densidade) de probabilidade de Z pode ser obtida a partir da função (de densidade) conjunta de probabilidade de (X,Y) e atendendo à função de transformação $\Phi(X,Y)$

- De forma análoga ao capítulo anterior podemos:
  - definir a covariância como o valor esperado de uma variável transformada
  - calcular directamente o valor esperado e a variância de variáveis transformadas

Slide 5.55

Transformação de pares de V.A. – Caso discreto

Nota: tópico não avaliado, apresentado apenas a título ilustrativo

Exemplo dos slides 5.3 e 5.4

Conhecidos o peso (X) e a altura (Y) dos elementos de uma população

V: Peso por unidade de altura

$V = X/Y$	$p_V(v)$
42.61 = 75/1.76	2/7
42.70 = 76/1.78	1/7
43.26 = 77/1.78	1/7
43.35 = 75/1.73	1/7
44.38 = 79/1.78	1/7
44.51 = 81/1.82	1/7

$$\mu_V = 42.61 \cdot \frac{2}{7} + 42.70 \cdot \frac{1}{7} + \ldots + 44.51 \cdot \frac{1}{7} = 43.346$$

$$\sigma_V^2 = (42.61 - 43.3)^2 \cdot \frac{2}{7} + \ldots + (44.51 - 43.3)^2 \cdot \frac{1}{7} = 0.562$$

- Tal como no caso unidimensional, os momentos populacionais da V.A. Z podem ser calculados a partir da função de probabilidade conjunta $p_Z(z)$ ou a partir de uma função de V.A. $Z = \Phi(X,Y)$

$$\mu_i = E(Z^i) = \sum_z z \cdot p_Z(z) = \sum_x \sum_y [\Phi(x,y)]^i \cdot p_{XY}(x,y)$$

$$\mu_i' = E\left[(Z - E(Z))^i\right] = \sum_z (z - E(Z))^i \cdot p_Z(z) = \sum_x \sum_y (\Phi(x,y) - E(Z))^i \cdot p_{XY}(x,y)$$

Slide 5.56

Transformação de pares de V.A. – Caso contínuo

Nota: tópico não avaliado, apresentado apenas a título ilustrativo

- Dada a V.A. (X,Y) com função de densidade de probabilidade conjunta $f_{XY}(x,y)$ e a função de transformação, tal que estas equações possam ser univocamente resolvidas em ordem a X e Y:

$$\begin{cases} Z = h(X,Y) \\ W = g(X,Y) \end{cases} \longrightarrow \begin{cases} X = G_1(Z,W) \\ Y = G_2(Z,W) \end{cases}$$

  Nota: Caso de $g(X,Y)$ não estar definida, podemos fazer $g(X,Y) = X$

- e existam e sejam contínuas as derivadas: $\frac{\delta x}{\delta z}$, $\frac{\delta x}{\delta w}$, $\frac{\delta y}{\delta z}$, $\frac{\delta y}{\delta w}$
- …pretende-se calcular $f_{ZW}(z,w)$

$$f_{ZW}(z,w) = \sum_i \left(f_{XY}(x,y) \cdot |J(z,w)|\right) \Bigg|_{\substack{x = G_1(z,w) \\ y = G_2(z,w)}}$$

- …em que:

$$J(z,w) = \begin{vmatrix} \frac{\delta x}{\delta z} & \frac{\delta x}{\delta w} \\ \frac{\delta y}{\delta z} & \frac{\delta y}{\delta w} \end{vmatrix} \neq 0$$

$\Longrightarrow$ Atenção ao *domínio* de $f_{ZW}(z,w)$

Slide 5.57

- E se se pretender apenas a distribuição marginal de Z (e.g. $h_Z(z)$)?

  Determina-se a função densidade de probabilidade marginal a partir da função de densidade de probabilidade conjunta

- E se $h(X,Y)$ e $g(X,Y)$ não forem biunívocas?

  Tal como no caso unidimensional divide-se o domínio em regiões onde tal aconteça

- Tal como no caso unidimensional, os momentos populacionais da V.A. Z podem ser calculados a partir da função de densidade de probabilidade conjunta $p_Z(z)$ ou a partir de uma função de V.A. $Z = \Phi(X,Y)$

$$\mu_i = E(Z^i) = \int_{\Re} z \cdot f_Z(z)\, dz = \int\int_{\Re^2} [\Phi(x,y)]^i \cdot f_{XY}(x,y)\, dy\, dx$$

$$\mu_i' = E\left[(Z - E(Z))^i\right] = \int_{\Re} (z - E(Z))^i \cdot p_Z(z) =$$

$$= \int\int_{\Re^2} (\Phi(x,y) - E(Z))^i \cdot f_{XY}(x,y)\, dy\, dx$$

Slide 5.58

4.9.6 Covariância definida como valor esperado de uma variável aleatória transformada

Definindo a variável transformada Z como:

$$Z = (X - E(X)) \cdot (Y - E(Y))$$

Caso discreto:

$$\mu_Z = E(Z) = \sum_x \sum_y (x - E(X)) \cdot (y - E(Y)) \cdot p_{XY}(x_i, y_j) = \gamma_{XY}$$

Caso contínuo:

$$\mu_Z = E(Z) = \int\int_{\Re^2} (x - E(X)) \cdot (y - E(Y)) \cdot f_{XY}(x_i, y_j)\, dy\, dx = \gamma_{XY}$$

Ou seja:

$$\gamma_{XY} = E\left[(x - E(X)) \cdot (y - E(Y))\right] = E(X \cdot Y - X \cdot E(Y) - Y \cdot E(X) + E(X) \cdot E(Y))$$
$$= E(X \cdot Y) - E(X \cdot E(Y)) - E(Y \cdot E(X)) + E(E(X) \cdot E(Y))) = E(XY) - E(X) \cdot E(Y)$$

Notas:

- expressão anterior já foi utilizada na resolução da alínea 8. do exemplo
- $E(X)$ e $E(Y)$ são constantes, logo $E(X \cdot E(Y)) = E(X) \cdot E(Y)$

Slide 5.59

4.9.7 Cálculo do valor esperado e da variância de pares variáveis transformadas

Tal como no caso unidimensional:

- começando por calcular a função (densidade) de probabilidade da variável transformada e calculando μ e σ^2 pelas suas definições (ver slides 5-57 e 5-58)
- ou calculando, de forma exacta ou aproximada, directamente a partir dos parâmetros da distribuição da variável original:

Combinação linear de variáveis (cálculo exacto)
Seja $Z = \Phi(X,Y) = a + bX + cY$, então

$$E(Z) = a + bE(X) + cE(Y)$$
$$V(Z) = b^2 V(X) + 2bc\,Cov(X,Y) + c^2 V(Y)$$

Notas:

1. $Var(X \pm Y) = Var(X) \pm 2Cov(X,Y) + Var(Y)$
2. Se X e Y são V.A. independentes, $Cov(X,Y) = 0$ vindo $Var(X - Y) = Var(X + Y) = Var(X) + Var(Y)$

Slide 5.60

Combinação não linear de variáveis (cálculo aproximado)
Nota: tópico não avaliado, apresentado apenas a título ilustrativo

- Neste caso $E(Z)$ e $V(Z)$ dependerão, no caso geral, não só de $E(X)$ e $E(Y)$, das suas variâncias e da covariância entre elas, mas também de momentos de ordem superior
- ... aproximação linear em torno de μ_X e μ_Y, por desenvolvimento em série de Taylor:

$$\Phi(X,Y) \approx \Phi(\mu_X, \mu_Y) + (X - \mu_X)\left[\frac{\delta\Phi}{\delta X}\right]_{\substack{X=\mu_X \\ Y=\mu_Y}} + (Y - \mu_Y)\left[\frac{\delta\Phi}{\delta Y}\right]_{\substack{X=\mu_X \\ Y=\mu_Y}}$$

Após alguma manipulação:

$$E(Z) \approx \Phi(E(X), E(Y))$$
$$V(Z) \approx \left[\frac{\delta\Phi}{\delta X}\right]^2_{\substack{X=\mu_X \\ Y=\mu_Y}} V(X) + \left[\frac{\delta\Phi}{\delta Y}\right]^2_{\substack{X=\mu_X \\ Y=\mu_Y}} V(Y)$$
$$+2\left[\frac{\delta\Phi}{\delta X}\right]_{\substack{X=\mu_X \\ Y=\mu_Y}} \left[\frac{\delta\Phi}{\delta Y}\right]_{\substack{X=\mu_X \\ Y=\mu_Y}} Cov(X,Y)$$

Slide 5.61

4.9.8 Exemplo – Livraria (cont.)

10. Determine a probabilidade de se venderem menos livros de inglês do que de francês ($P(X<Y)$)

$$P(X<Y)=\sum_y\sum_{x<y}p_{XY}(x,y)=p_{XY}(0,1)+p_{XY}(0,2)+\ldots+p_{XY}(2,3)=$$
$$=0.04+0.03+0.09+0.01+0.06+0.03=0.26=26\%$$

11. Determine o valor esperado do número de livros de língua estrangeira vendidos semanalmente, $E(X+Y)$

$$E(X+Y)=\sum_x\sum_y(x+y)p_{XY}(x,y)=0\cdot p_{XY}(0,0)+\ldots+7\cdot p_{XY}(4,3)=$$
$$=0\cdot 0.02+1\cdot 0.04+2\cdot 0.03+\ldots+7\cdot 0.02=3.48$$

Alternativamente: $E(X+Y)=E(X)+E(Y)=1.9+1.58=3.48$

Slide 5.62

12. Considere a V.A. Z obtida pelo produto das V.A. X e Y ($Z=X\cdot Y$). Obtenha a função de probabilidade de Z ($p_Z(z)$)

z	0	1	2	3	4	6	8	9	12
$p_Z(z)$	0.18	0.12	0.21	0.12	0.14	0.13	0.06	0.02	0.02

13. Calcule o valor esperado de Z (μ_Z)

$$\mu_Z=E(Z)=\sum_z z\cdot p_Z(z)=0\times 0.18+1\times 0.12+\cdots+12\times 0.02=3.14$$

Alternativamente e como conhecemos a função de probabilidade conjunta de X,Y podemos também calcular $E(Z)$ a partir de $E(XY)$

$$\mu_Z=E(Z)=E(XY)=3.14 \qquad \text{(já calculado anteriormente)}$$

14. Calcule a variância de Z (σ_Z^2)

$$\mu_{Z^2}=E(Z^2)=\sum_z z^2\cdot p_Z(z)=0^2\times 0.18+1^2\times 0.12+\cdots+12^2\times 0.02=17.3$$
$$\sigma_Z^2=Var(Z)=E(Z^2)-E^2(Z)=17.3-3.14^2=7.44$$

Slide 5.63

15. Calcule a variância de $X-Y$ (σ_{X-Y}^2)

$$\sigma_{X-Y}^2=Var(X-Y)=Var(X)+Var(Y)-2Cov(X,Y)=$$
$$=1.29+1.58-2\times 0.138=2.594$$

16. ~~Suponha agora que se conhecem apenas os parâmetros das V.A. X e Y ($\mu_X=1.9$, $\mu_Y=1.58$, $\sigma_X^2=1.29$ e $\sigma_Y^2=1.58$), assim como conhecemos a covariância entre X e Y ($\gamma_{XY}=0.138$). Ou seja, desconhecemos $p_{XY}(x,y)$ assim como $p_Z(z)$.~~
~~Calcule μ_Z e μ_Z com $Z=XY$.~~

$$\mu_Z=E(Z)\approx E(X)\cdot E(Y)=1.9\times 1.58=3.002$$
$$\sigma_Z^2=Var(Z)\approx E^2(Y)\cdot Var(X)+E^2(X)\cdot Var(Y)+2\cdot E(Y)\cdot E(X)\cdot Cov(X,Y)=$$
$$=1.58^2\times 1.29+1.9^2\times 0.7236+2\times 1.58\times 1.29\times 0.138=6.395$$

(comparar estes valores aproximados com os exactos calculados antes) Slide 5.64

4.10 Exercícios

1. Três baterias eléctricas são escolhidas aleatoriamente de um grupo contendo 3 novas, 4 usadas e 5 avariadas. Considere que as baterias novas funcionam a 100% da sua capacidade, as usadas a 80% e as avariadas a 0%. Cada bateria debita 10 V.

i) Sendo X e Y respectivamente o número de baterias novas e usadas, verifique se as variáveis são independentes.

Cálculo de $p_{XY}(x,y)$:

$p(0,0)=\binom{5}{3}/\binom{12}{3}=10/220$
$p(0,1)=\binom{4}{1}\binom{5}{2}/\binom{12}{3}=40/220$
$p(0,2)=\binom{4}{2}\binom{5}{1}/\binom{12}{3}=30/220$
$p(0,3)=\binom{4}{3}/\binom{12}{3}=4/220$
$\ldots$

	Y				
X	0	1	2	3	$p_X(x)$
0	0.045	0.182	0.136	0.018	0.381
1	0.136	0.273	0.082	0	0.491
2	0.068	0.055	0	0	0.123
3	0.005	0.00	0	0	0.005
$p_Y(y)$	0.254	0.51	0.218	0.018	1

São dependentes porque $P(X=3\cap Y=3)=0 \neq P(X=3)\cdot P(Y=3)$

ii) Calcule o valor esperado e a variância da variável Voltagem total (V_T).

$$\mu_{V_T} = E(V_T) = \sum_X \sum_Y (X \cdot 10 \cdot 1 + Y \cdot 10 \cdot 0.8) \cdot p_{XY}(x,y) = 15.5$$

$$\sigma^2_{V_T} = Var(V_T) = \sum_X \sum_Y [(X \cdot 10 \cdot 1 + Y \cdot 10 \cdot 0.8) - E(V_T)]^2 \cdot$$
$$\cdot p_{XY}(x,y) = 48.2$$

Slide 5.65

2. Tendo em conta o historial dos confrontos entre duas equipas de futebol nos últimos 40 anos, definiu-se a função de probabilidade para os resultados dos jogos entre as duas equipas.

# de golos marcados	Equipa A: 0	1	2	$p_B(b)$
Equipa B: 0	0.06	0.08	0.06	0.2
1	0.14	0.1	0.18	0.42
2	0.07	0.2	0.06	0.33
3	0.03	0.02	0	0.05
$p_A(a)$	0.3	0.4	0.3	1

Política de prémios da equipa A	
Vitória:	50000$
Empate:	5000$
Derrota:	-2000$
Não sofrer golos:	10000$

i) Verifique se o número de golos marcados pela equipa B é independente do número de golos marcados pela equipa A.

$$P(A = 2 \cap B = 3) = 0 \quad \neq \quad P(A = 2) \cdot P(B = 3) = 0.3 \times 0.05 = 0.015$$

ii) Calcule a probabilidade de a equipa B vencer o próximo confronto (G) se marcar pelo menos um golo (M).

$$P(G|M \geq 1) = \frac{P(G \cap M \geq 1)}{P(M \geq 1)} = \frac{0.14 + 0.07 + 0.2 + 0.03 + 0.02}{1 - 0.2} = 0.575$$

iii) Calcule o valor esperado do prémio (P) recebido por cada jogador da equipa A no próximo confronto entre as duas equipas.

$$E(P) = 50 \cdot (0.08 + 0.06 + 0.18) + 5 \cdot (0.06 + 0.1 + 0.06) -$$
$$-2 \cdot (0.14 + 0.07 + 0.2 + 0.03 + 0.02) + 10 \cdot 0.2 = 18.18$$

Slide 5.66

3. Duas tarefas (X e Y) têm de ser executadas pelo mesmo operário.

   X : número de minutos para realizar a tarefa 1; $\mu_X = 20$, $\sigma_X = 5$

   Y : número de minutos para realizar a tarefa 2; $\mu_Y = 30$, $\sigma_Y = 8$

   Sabe-se que X e Y são independentes.

   Qual é a média e o desvio padrão do tempo necessário para executar as duas tarefas (W)?

$$W = X + Y$$

$$\mu_W = E(X + Y) = E(X) + E(Y) = \mu_X + \mu_Y = 20 + 30 = 50$$

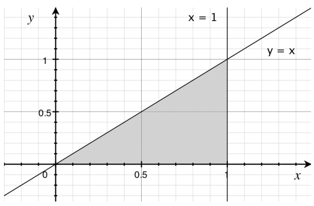
$$\sigma^2_W = Var(X + Y) = Var(X) + Var(Y) + 2 \cdot Cov(X,Y) =$$
$$= \sigma^2_X + \sigma^2_Y + 0 = 5^2 + 8^2 = 89$$

$$\sigma_W = \sqrt{\sigma^2_W} = \sqrt{89} = 9.434$$

Slide 5.67

4. Seja (X,Y) uma variável aleatória bidimensional contínua com a seguinte função densidade de probabilidade conjunta:

$$f_{XY}(x,y) = \begin{cases} 2, & 0 < x < 1 \wedge 0 < y < x \\ 0, & \text{outros valores} \end{cases}$$



i) Verifique se as variáveis aleatórias X e Y são independentes.
Cálculo das funções densidade marginais:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy = \int_0^x 2\, dy = 2x, \text{ para } 0 < x < 1$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dx = \int_y^1 2\, dx = 2\,(1-y), \text{ para } 0 < y < 1$$

Como $f_{XY}(x,y) \neq f_X(x) \cdot f_Y(y)$ as variáveis X e Y não são independentes.

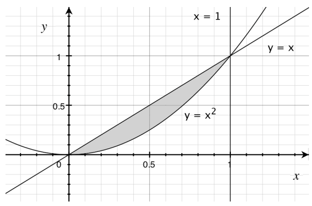
ii) Calcule $E(Y|X)$.

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{2}{2x} = \frac{1}{x} \text{ , para } 0 < y < x \wedge 0 < x < 1$$

$$E(Y|X) = \int_{-\infty}^{+\infty} y\, f_{Y|X}(y|x)\, dy = \int_0^x y\, \frac{1}{x}\, dy = \frac{x}{2} \text{ , para } 0 < x < 1$$

Slide 5.68

iii) Calcule $P(Y > X^2)$.



$$\mathrm{P}\left(Y > X^2\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy\, dx = \int_0^1\int_{x^2}^{x} 2\, dy\, dx = 2\int_0^1 [y]_{x^2}^{x}\, dx =$$

$$= 2\int_0^1 (x - x^2)\, dx = 2\,\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = 2 \times \frac{1}{6} = \frac{1}{3}$$

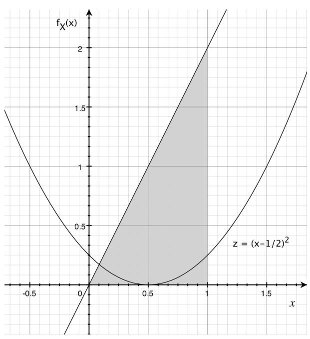
iv) ~~Considere uma nova variável aleatória $Z = (X - \frac{1}{2})^2$. Determine a função densidade de probabilidade de Z, $f_Z(z)$.~~

$$f_Z(z) = \left.\frac{f_X(x)}{\left|\frac{dz}{dx}\right|}\right|_{X=H^{-1}(z)} = \left. f_X(x) \cdot \left|\frac{dx}{dz}\right|\right|_{X=H^{-1}(z)}$$

Slide 5.69

$$z = \left(x - \frac{1}{2}\right)^2 \Leftrightarrow x - \frac{1}{2} = \pm\sqrt{z} \Leftrightarrow x = \frac{1}{2} \pm \sqrt{z}$$

$$\left|\frac{dx}{dz}\right| = \left|\frac{1}{2\,\sqrt{z}}\right| = \frac{1}{2\,\sqrt{z}}$$



$$f_Z(z) = f_X(x)\,\left|\frac{dx}{dz}\right|_{x=\frac{1}{2}-\sqrt{z}} + f_X(x)\,\left|\frac{dx}{dz}\right|_{x=\frac{1}{2}+\sqrt{z}}$$

$$= 2\,\left(\frac{1}{2} - \sqrt{z}\right) \times \frac{1}{2\,\sqrt{z}} + 2\,\left(\frac{1}{2} + \sqrt{z}\right) \times \frac{1}{2\,\sqrt{z}}$$

$$= \frac{1}{2\,\sqrt{z}} - 1 + \frac{1}{2\,\sqrt{z}} + 1 = \frac{1}{\sqrt{z}}$$

$$0 < x < 1 \qquad \Longrightarrow \qquad 0 < z < \frac{1}{4}$$

Ou seja:

$$f_Z(z) = \begin{cases} \frac{1}{\sqrt{z}} & , 0 < z < \frac{1}{4} \\ 0 & , \text{outros valores} \end{cases}$$

Slide 5.70

4.11 Anexos

4.11.1 Função geradora de momentos

Seja X uma v.a. qualquer (discreta ou contínua). Se existir um número positivo h tal que se possa definir a função

$$G(t) = E[e^{tX}]$$

para qualquer $t \in]-h,h[$, tal função designa-se por *função geradora de momentos* da variável X.

Propriedade importante:

$$\frac{d^nG}{dt^n}(0) = G^{(n)}(0) = E(X^n) \qquad (\text{com } n \geq 1)$$

Por exemplo:

$$G'(0) = E(X)$$
$$G''(0) = E(X^2)$$

Slide 5.71

4.11.2 Desigualdades de Markov e de Chebychev

- E se em vez de conhecermos a distribuição de probabilidade de uma v.a. conhecermos apenas os seus parâmetros?
- Seja X uma v.a. com $E(X) = \mu$ e $Var(X) = \sigma^2$ mas com função densidade de probabilidade $f(x)$ desconhecida. Então:

Desigualdade de Markov

$$P(X \geq C) \leq \frac{\mu}{C}, \quad \text{se } X \text{ for uma v.a. não negativa e com } C > 0$$

Permite saber a probabilidade da cauda de uma distribuição conhecendo apenas o seu valor esperado

Desigualdade de Chebychev

$$P[|X-\mu| \geq C\sigma] \leq \frac{1}{C^2}, \quad \forall_{C>0}$$

Limite superior para a probabilidade de uma dada v.a tomar valores fora de um certo intervalo, centrado na média da v.a.

Slide 5.72