Análise de Risco

Sessão 6
Modelagem da corrrelação

Prof. E.A. Schmitz

PPGI/UFRJ

July 23, 2018

Revisão - 5 teoremas importantes

- T1- E(a+bX) = a + bE(X).
- T2- E(X + Y) = E(X)+E(Y).
- T3- E(XY) = E(X)·E(Y) indep X e Y.
- T4- V(X+Y)=V(X)+V(Y)+ 2*Cov(X,Y)
  - Onde $Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$
- T5- V(X + Y) = V(X) + V(Y) se indep X e Y, Cov(X,Y)=0

A omissão da modelagem da correlação pode afetar muito a análise de risco de um projeto. Por quê?

Definições

1-Função de probabilidade conjunta
$f_{xy}(x,y) = Prob(X = x \text{ e } Y = y)$

2-Função de probabilidade marginal
$f_x(x,y) = \sum_y f_{xy}(x,y)$
$f_y(x,y) = \sum_x f_{xy}(x,y)$

3-Sejam duas VAs X,Y, com função de probabilidade conjunta f_{xy} e marginais $f_x e f_y$ e respectivamente. Dizemos que X e Y são independentes sse:
$f_{xy}(x,y) = f_x(x) * f_y(y)$

Exemplo

A tabela abaixo mostra distribuição conjunta de probabilidade dos pontos obtidos ao jogar dois dados de 4 faces.

- calcule as duas funções de probabilidades marginais
- avalie se as variáveis são independentes

4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0
	1	2	3	4

Dependência e risco 1

Variáveis do modelo podem apresentar alguma forma de dependência entre si.
1-Observe este fragmento de modelo em R

- $c1 < -rtriangle(1000, 10, 30, 20)$
- $c2 < -rtriangle(1000, 20, 60, 40)$
- $plot(c1, c2)$

2-repita com:

- $e1 < -rnorm(1000, 0, 5)$
- $c3 < -2 * c1 + 10 + e1$
- $plot(c1, c3)$

3-repita com:

- $c4 < -80 - 2 * c1 + e1$
- $plot(c1, c4)$

Dependência e risco 2

4-Observe este fragmento de modelo em R

- $cov(c1, c2)$
- $m1 < -mean(c1)$
- $m2 < -mean(c2)$
- $sp < -(c1 - m1) * (c2 - m2)$
- $head(sp)$

5-repita com:

- $library(MASS)$
- $cormat < -Sigma = matrix(c(1, 1, 1, 1), ncol = 2)$
- $cc < -mvrnorm(500, mu = c(0, 0), Sigma = cormat)$
- $plot(cc[, 1], cc[, 2])$

Coeficientes de correlação de Pearson

Coeficiente de correlação (Pearson):
$Cor(X, Y) = Cov(X, Y)/(\delta_x * \delta_y)$
O coeficiente de correlação é a covariancia normalizada.
O valor de Cor(.,.) varia entre (-1..0..1).
Se Cor (X,Y)=

- 0: não há correlação pois a covariancia é 0.
- +1:forte correlação positiva.
- -1: forte correlação negativa.

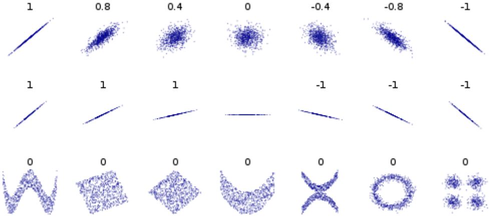
Coeficientes de correlação de Spearman

Coeficiente de correlação de Spearman ρ
Estatística não paramétrica para indicar a relação entre 2 VAs
Algoritmo:

- 1- encontre a ordem das observações de X e Y (u_i, v_i), onde i=1 corresponde ao maior valor.
- 2- calcule $\rho = 1 - 6 * \sum \frac{(u_i - v_i)^2}{(n(n^2-1))}$

Exemplos

Diversos tipos de correlação entre variáveis aleatórias



Funções para correlação e covariância em R

cov(x, y = NULL, use = "everything", method = c("pearson", "kendall", "spearman"))

cor(x, y = NULL, use = "everything", method = c("pearson", "kendall", "spearman"))

Como gerar amostras correlacionadas

Técnicas mais utilizadas

- 1-usando uma matriz de correlação
- 2-técnica de look-up
- 3-técnica da envoltória (envelope method)

Geração de amostras normais correlacionadas em R

- 1-instalar o pacote MASS
- 2-usar a função mvrnorm(m,mu,Sigma): retorna uma lista de m amostras de (n) variáveis aleatórias normalmente distribuidas.
  - m: número de amostras
  - mu: vetor das médias das n variáveis
  - Sigma: matriz de correlação entre as n variáveis

Exemplo:

x< −mvrnorm(500,mu=c(0,0),Sigma=matrix(c(1,1,1,1),ncol=2))

Geração de amostras triangulares correlacionadas em R

Modelo: n VAs distintas (v1,v2,..vn), interligadas por uma matriz de correlação A [n,n] Método

- 1-gerar Z[m,n]: m n-plas amostras normais reduzidas correlacionadas por A
  Z¡-mvrnorm(m,mu=rep(0,times=n),Sigma=A)
- 2-gerar U com m n-plas uniformes U¡-pnorm(Z)
- 3-gerar matriz de triangulares correlacionadas
  T[,1]¡-qtriangle(U[,1],min1,max1,mp1)

Geração de amostras correlacionadas (não-paramétrica)

Algoritmo não-paramétrico de Iman&Conover
Entradas:
$X[N, k]$: matriz de amostras não correlacionadas
C:matriz de correlação de Spearman entre as k variáveis
Saida:
Y[N,k]: matriz de amostras correlacionas por C
Passos

1. ache $P.P^T = C$
2. $R \leftarrow matrix[N, k]$
3. $S < -\phi^{-1}(i/N + 1)\forall i \in \{1..N\}$
4. $R[, j] \leftarrow perm(S)\forall j \in \{1..k\}$
5. $R^* \leftarrow R.P^T$
6. $Y[, j] < -X[, j]$ ordered by $R^*[, j]$