Probabilidad

Daniel Fraiman

Maestría en Ciencia de Datos, Universidad de San Andrés

UN RECORRIDO A VUELO DE PÁJARO POR VECTORES ALEATORIOS

Vectores aleatorios

Vamos a estudiar

- vectores aleatorios discretos.
- vectores aleatorios continuos.

¿Qué hay de nuevo?

- La dependencia entre las coordenadas (o variables) del vector.

Vectores aleatorios

Ejemplos

Elijo una persona al azar de la población y

- le mido (altura,peso) [(X, Y) continuas]
- le “mido” (número de hijos, número de casamientos) [(X, Y) discreta]

Vectores aleatorios en 2D

(X, Y) vector aleatorio continuo (X e Y continuas).
(X, Y) vector aleatorio discreto (X e Y discretos).

	vector discreto	vector continuo
Conjunta	$\mathbb{P}(X = x, Y = y)$	$f_{X,Y}(x, y)$
Marginal	$\mathbb{P}(X = x)$ & $\mathbb{P}(Y = y)$	$f_X(x)$ & $f_Y(y)$
Condicional	$\mathbb{P}(X = x \mid Y = y)$ & $\mathbb{P}(Y = y \mid X = x)$	$f_{X\mid Y=y}(x)$ & $f_{Y\mid X=x}(y)$

Si la persona tiene una altura de 1.75m, cual seria la densidad del peso
Suponiendo que X e Y son peso y altura

Vectores aleatorios en 2D

	la conjunta verifica
vector discreto	$\sum_{i,j} \mathbb{P}(X = x_i, Y = y_j) = 1$
vector continuo	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$

	definición de la marginal
vector discreto	$\mathbb{P}(X = x_i) = \sum_j \mathbb{P}(X = x_i, Y = y_j)$
vector continuo	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$

Calculo de las probabilidades marginales a partir de las prob. conjuntas

	definición de la condicional
vector discreto	$\mathbb{P}(X = x \mid Y = y) = \frac{\mathbb{P}(X=x,Y=y)}{\mathbb{P}(Y=y)}$
vector continuo	$f_{X\mid Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

Calculo de las probabilidades condicionales a partir de las prob. conjunta y marginal

Vectores aleatorios en 2D

(X, Y) vector aleatorio continuo (X e Y continuas).
(X, Y) vector aleatorio discreto (X e Y discretos).

	vector discreto	vector continuo
Conjunta	$\mathbb{P}(X = x, Y = y)$	$f_{X,Y}(x, y)$
Marginal	$\mathbb{P}(X = x)$ & $\mathbb{P}(Y = y)$	$f_X(x)$ & $f_Y(y)$
Condicional	$\mathbb{P}(X = x \mid Y = y)$ & $\mathbb{P}(Y = y \mid X = x)$	$f_{X\mid Y=y}(x)$ & $f_{Y\mid X=x}(y)$

Esperanza matematica

	vector discreto	vector continuo
$\mathbb{E}((X, Y))$	$(\mathbb{E}(X), \mathbb{E}(Y))$	$(\mathbb{E}(X), \mathbb{E}(Y))$
$\mathbb{E}(X)$	$\sum_i x_i \mathbb{P}(X = x_i)$	$\int_{-\infty}^{\infty} x f_X(x) dx$
$\mathbb{E}(X \mid Y = y)$	$\sum_i x_i \mathbb{P}(X = x_i \mid Y = y)$	$\int_{-\infty}^{\infty} x f_{X\mid Y=y}(x) dx$

Esperanza condicional—>”en promedio cuanto pesa una persona, si yo se que mide tanto”

Suma de los valores multiplicado por la probabilidad condicional (vector discreto)
Integral de los valores, multiplicado por la probabilidad condicional(vector continuo)

Independencia y Covarianza

Independencia

Dos variables X e Y son independientes si

$$\mathbb{P}(X \leq a, Y \leq b) = \mathbb{P}(X \leq a)\,\mathbb{P}(Y \leq b) \qquad \forall a, b \in \mathbb{R}.$$

Definicion de INDEPENDENCIA

Independencia y Covarianza

$$
\begin{aligned}
COV(X,Y) &= \mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right) \\
&= \begin{cases} \sum\limits_{i,j}(x_i-\mathbb{E}(X))(y_j-\mathbb{E}(Y))\mathbb{P}(X=x_i,Y=y_j) & \text{discr.} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mathbb{E}(X))(y-\mathbb{E}(Y))f_{X,Y}(x,y)dxdy & \text{contin.} \end{cases}
\end{aligned}
$$

La covarianza, solo mide el grado de dependencia LINEAL, con lo cual una covarianza de 0, puede ocurrir en un caso donde haya una gran dependencia de los datos, pero que la misma NO SEA LINEAL, y por eso no la puede captar la covarianaza.

Covarianza

Propiedades

1. $COV(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
2. $COV(X,Y) = COV(Y,X)$
3. $COV(X,X) = Var(X)$
4. $COV(aX,bY) = a \cdot b \cdot COV(X,Y)$
5. $COV(X+a,Y+b) = COV(X,Y)$
6. $COV(X+Y,Z+W) =$ $COV(X,Z)+COV(X,W)+COV(Y,Z)+COV(Y,W)$ [4 variables]
7. $Var(X+Y) = Var(X)+Var(Y)+2COV(X,Y)$
8. Si X e Y son independientes entonces $COV(X,Y) = 0$

Covarianza

La versión normalizada de la covarianza es el coeficiente de correlación lineal,

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

1. ρ mide el grado de dependencia lineal entre X e Y
2. $-1 \leq \rho \leq 1$
3. X e Y caen en un recta si $|\rho| = 1$

Covarianza

Σ= matriz de covarianza

$$\Sigma = \begin{pmatrix} COV(X,X) & COV(X,Y) \\ COV(Y,X) & COV(Y,Y) \end{pmatrix} =$$
$$\begin{pmatrix} Var(X) & COV(X,Y) \\ COV(X,Y) & Var(Y) \end{pmatrix} =$$
$$\begin{pmatrix} Var(X) & \rho\sqrt{Var(X)Var(Y)} \\ \rho\sqrt{Var(X)Var(Y)} & Var(Y) \end{pmatrix}$$

Si tenemos un vector en dimensión 4 (X,Y,Z,W)

$$\Sigma = \begin{pmatrix} Var(X) & COV(X,Y) & COV(X,Z) & COV(X,W) \\ COV(X,Y) & Var(Y) & COV(Y,Z) & COV(Y,W) \\ COV(X,Z) & COV(Y,Z) & Var(Z) & COV(Z,W) \\ COV(X,W) & COV(Y,W) & COV(Z,W) & Var(W) \end{pmatrix}$$

Vectores aleatorios muy utilizados

Vectores aleatorios continuos muy utilizados

- Normal Multivariada

Vectores aleatorios discretos muy utilizados

- Multinomial (generalización de la Binomial)

Normal Multivariada

$$(X,Y) \sim N((0,0),\Sigma)$$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}-2\frac{\rho xy}{\sigma_x\sigma_y}\right)}$$

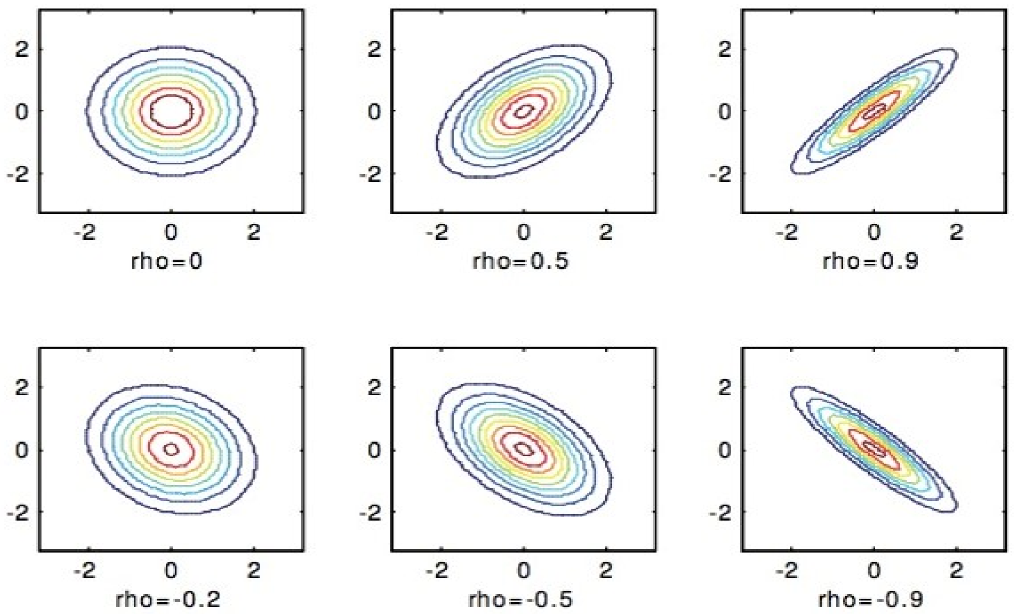
$$\vec{X} \sim N(\vec{0},\Sigma)$$

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}\vec{x}^T\Sigma^{-1}\vec{x}}$$

$$\vec{X} \sim N(\vec{\mu},\Sigma)$$

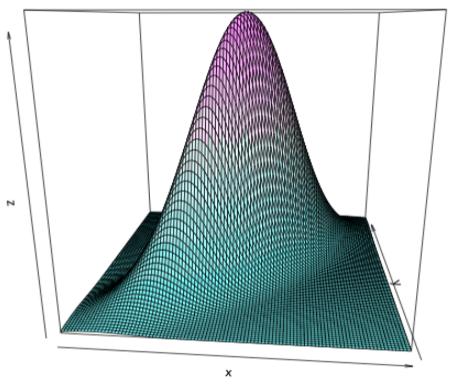
$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})}$$

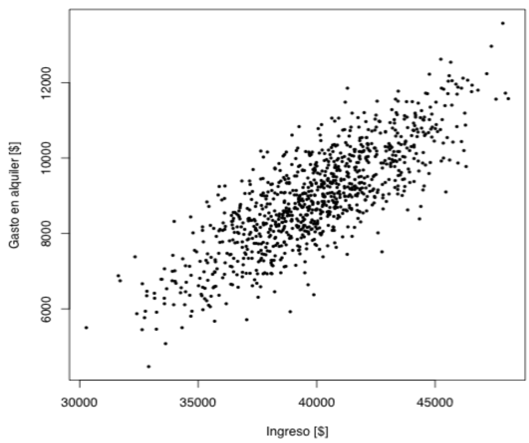
Normal Multivariada



rho(coef de correlacion lineal) me dice el nivel de correlacion lineal que tiene, a mayor rho, mayor correlacion y mejor nivel de prediccion lineal

Normal Multivariada





Normal Multivariada

matriz de covarianza

Pregunta

- $(X, Y) \sim N(\vec{\mu}, \Sigma)$, ¿Cómo tiene que ser Σ para que X e Y sean indep?
- ¿Cómo serán las distribuciones marginales de una Normal Multivariada? **quedaria una matriz diagonal**
- ¿Cómo serán las distribuciones condicionales de una Normal Multivariada? **quedaria una distribucion normal**

Multinomial

Ahora cada categoria (A, B, C,...,), tendra su propia probabilidad de exito, pero PA+PB+PC+...+Pn = 1

$\vec{X} \sim Multinomial(n, \vec{p})$

Ahora tenemos K categorías (en vez de 2) y realizamos n experimentos independientes.

Llamemos X_i= número de veces que salió la categoría i en n exp.

$$\mathbb{P}\left(\vec{X} = (x_1, x_2, \ldots, x_K)\right) = \frac{n!}{x_1!x_2!\ldots x_K!} p_1^{x_1} p_2^{x_2} \ldots p_K^{x_K}$$

con $\sum_{i=1}^{K} x_i = n$ y $\sum_{i=1}^{K} p_i = 1$.

Multinomial

$\vec{X} \sim Multinomial(n, \vec{p})$

$$\mathbb{P}\left(\vec{X} = (x_1, x_2, \ldots, x_K)\right) = \frac{n!}{x_1!x_2!\ldots x_K!} p_1^{x_1} p_2^{x_2} \ldots p_K^{x_K}$$

con $\sum_{i=1}^{K} x_i = n$ y $\sum_{i=1}^{K} p_i = 1$.

Preguntas

- ¿ X_i y X_j son independientes?
- ¿ Cuál será la distribución marginal?
- ¿ Cuál será la distribución condicional?

Xi y Xj, me dicen cuantos experimentos cayeron en cada categoria—> Son DEPENDIENTES, ya que la suma de todas las categorias debe ser n

La distribución marginal, seria hacer una binomial. pd(si estudiara la categoria "d"), y n seria el total.

n =100 Xmultinomial(100, (p1, p2, p3, p4))

P(x1,x2,x3,x4) = (…))
P((x1,x2,x3) = (x1, x2,x3)I(x4 = 5)) —> Y=(x1,x2,x3) multinom(95, (p1, p2, p3)/p1+p2+p3))