

Teoría de Probabilidad - Variable aleatoria Continua

Cristian Guarnizo-Lemus
cristianguarnizo@itm.edu.co

Maestria en Automatización y Control Industrial

Contenido

1 Variables aleatorias continuas

- Funciones de densidad de probabilidad
- Ejemplos de funciones de densidad de probabilidad

2 Vectores aleatorios

- Preliminares
- fdp Gaussiana multivariada

3 Momentos estadísticos

Funciones de densidad de probabilidad

- Como se mencionó anteriormente, una *variable aleatoria continua* puede tomar valores en un intervalo de la recta real.
- La ley de probabilidad para una variable aleatoria continua X se define en términos de la función de densidad de probabilidad (pdf) $f_X(x)$,

$$f_X(x) = \frac{d F_X(x)}{d x}.$$

Propiedades de una fdp

1 $f_X(x) \geq 0.$

2 $\int_{-\infty}^{\infty} f_X(x) dx = 1.$

3 $P(X \leq a) = F_X(a) = \int_{-\infty}^a f_X(x) dx.$

4 $P(a \leq X \leq b) = \int_a^b f_X(x) dx = F_X(b) - F_X(a).$

Propiedades de una fdp - Ejemplo

La distribución logística tiene una función distribución

$$F_X(x) = \frac{e^x}{1 + e^x}, \quad x \in \mathbb{R}$$

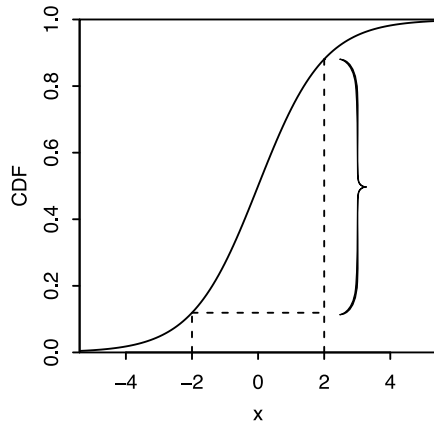
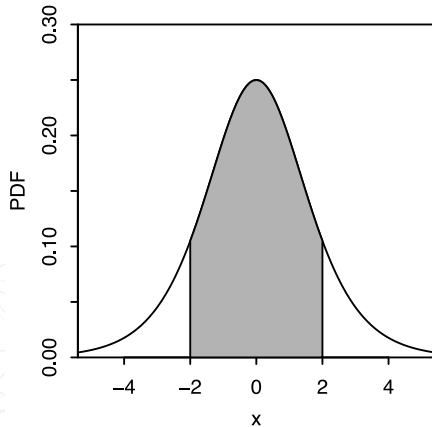
Se puede demostrar que su fdp esta dada por

$$f_X(x) = \frac{e^x}{(1 + e^x)^2}$$

Por ejemplo para calcular

$$P(-2 < X < 2) = \int_{-2}^2 \frac{e^x}{(1 + e^x)^2} dx = F(2) - F(-2) \approx 0.76.$$

Propiedades de una fdp - Ejemplo



Dos variables aleatorias: pdf conjunta

- Si existen dos variables aleatorias X y Y , pueden caracterizarse con la pdf $f_{X,Y}(x, y)$.
- Si la función de distribución conjunta $F_{X,Y}(x, y)$ es continua y tiene derivadas parciales, la pdf conjunta se define como

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}.$$

- Igualmente se tiene los siguientes resultados

1 $F_{X,Y}(x, y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(\alpha, \beta) d\alpha d\beta.$

2 $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(\alpha, \beta) = 1.$

Fdp marginal y fdp condicional

- Las funciones de densidad de probabilidad marginal están dadas como

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx.$$

- Las funciones de densidad de probabilidad condicional se definen como

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

$$f_{Y|X}(y|x) \triangleq \frac{f_{X,Y}(x, y)}{f_X(x)}.$$

Fdp marginal y fdp condicional - Ejercicio

Suponga la distribución fdp conjunta de las variables aleatorias X y Y

$$f_{X,Y}(x,y) = \begin{cases} 2(1-x), & 0 \leq x \leq 1 \text{ y } 0 \leq y \leq 1 \\ 0, & \text{en las demás regiones} \end{cases}$$

Calcular la probabilidad del evento $\{x \leq 0.5, 0.4 \leq y \leq 0.7\}$ y encontrar las fdp marginales de X y Y .

Teorema de Bayes e Independencia Estadística

- El teorema de Bayes para variables aleatorias continuas está dado como

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)f_Y(y)}{\int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(\lambda) d\lambda}.$$

- Se dice que dos variables aleatorias X y Y son estadísticamente independientes si

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

Valores esperados

Los valores esperados de variables aleatorias continuas se definen como

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

$$\mu_X = E\{X\} = \int_{-\infty}^{\infty} x f_X(x) dx,$$

$$\sigma_X^2 = E\{(X - \mu_X)^2\} = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx,$$

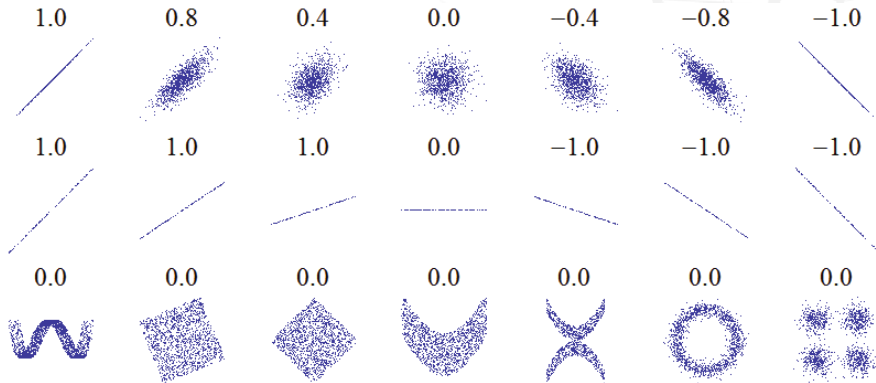
$$\sigma_{XY} = E\{(X - \mu_X)(Y - \mu_Y)\},$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y},$$

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

$$E\{g(X)h(Y)\} = E\{g(X)\}E\{h(Y)\}.$$

Valores esperados - Correlación



Contenido

1 Variables aleatorias continuas

- Funciones de densidad de probabilidad
- Ejemplos de funciones de densidad de probabilidad

2 Vectores aleatorios

- Preliminares
- fdp Gaussiana multivariada

3 Momentos estadísticos

Fdp uniforme

Se dice que una variable aleatoria continua X sigue una función de densidad de probabilidad uniforme si

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{de otra manera} \end{cases}$$

Cual es la función de distribución? Calcular media y varianza.

Fdp Gaussiana I

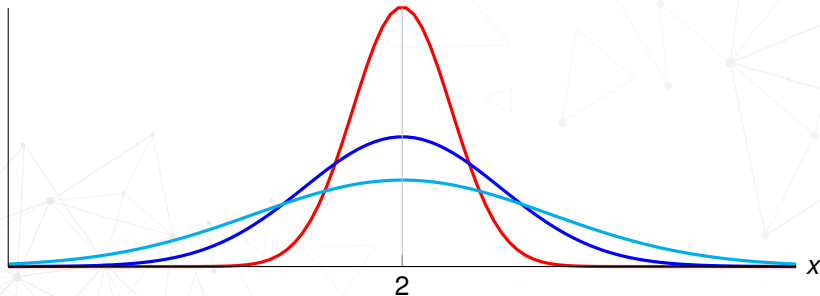
- La fdp Gaussiana está motivada en el *teorema del límite central*: una variable aleatoria determinada como la suma de un gran número de causas independientes tiende a tener una fdp Gaussiana.
- Bajo la presunción del teorema del límite central, la fdp Gaussiana se emplea para modelar el ruido eléctrico.
- La fdp Gaussiana tiene la forma

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp \left\{ -\frac{(x - \mu_X)^2}{2\sigma_X^2} \right\}.$$

- La familia de fdp Gaussianas está caracterizada sólo por dos parámetros μ_X y σ_X^2 , que son la media y la varianza de la variable aleatoria X .
- En este curso, la fdp Gaussiana se denota como $f_X(x) = \mathcal{N}(x|\mu_X, \sigma_X^2)$ ó $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$.

Fdp Gaussiana II

La media de las tres Gaussianas es $\mu_X = 2$, y las desviaciones estándar son $\sigma_X = 0.5$ (en rojo), $\sigma_X = 1$ (en azul) y $\sigma_X = 1.5$ (en cyan).



Fdp Gaussiana III

- En algunas aplicaciones resulta necesario calcular la probabilidad

$$P(X > a) = \int_a^{\infty} \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp \left\{ -\frac{(x - \mu_X)^2}{2\sigma_X^2} \right\} dx.$$

- Haciendo un cambio de variable $z = (x - \mu_X)/\sigma_X$, la integral anterior se reduce a

$$P(X > a) = \int_{(a-\mu_X)/\sigma_X}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z^2}{2} \right\} dz.$$

Fdp Gaussiana III

- Esta integral no puede resolverse analíticamente. Se expresa en términos de la función $Q(y)$ definida como

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^{\infty} \exp \left\{ -\frac{z^2}{2} \right\} dz, \quad y > 0.$$

La función $Q(y)$ se puede calcular a través de la relación

$$Q(y) = \frac{1}{2} \operatorname{erfc} \left(\frac{y}{\sqrt{2}} \right),$$

donde $\operatorname{erfc}(x)$ es la función de error complementaria.

Fdp Gaussiana IV

- Esta integral no puede resolverse analíticamente. Se expresa en términos de la función $Q(y)$ definida como

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^{\infty} \exp \left\{ -\frac{z^2}{2} \right\} dz, \quad y > 0.$$

La función $Q(y)$ se puede calcular a través de la relación

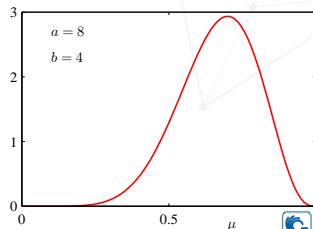
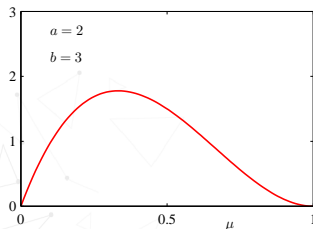
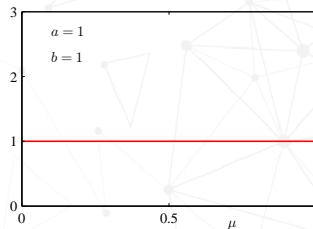
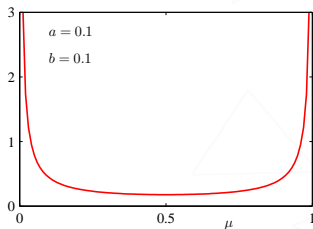
$$Q(y) = \frac{1}{2} \operatorname{erfc} \left(\frac{y}{\sqrt{2}} \right),$$

donde $\operatorname{erfc}(x)$ es la función de error complementaria.

Otras funciones de densidad de probabilidad continua

- **Gamma** (prior conjugado para la varianza de una Gaussiana o la media de una Poisson).
- **Beta** (prior conjugada para los parámetros de una binomial).
- **Dirichlet** (prior conjugada para los parámetros de una multinomial).
- **Exponencial** (se usa para modelar el tiempo de espera de un próximo evento en un proceso de Poisson)
- **Weibull** (se usa para modelar tiempos de fallas en análisis de confiabilidad).

Función de densidad de probabilidad Beta(μ, a, b)



Ejemplo: Beta-Bernoulli

Asuma que tiene datos $\mathcal{D} = \{x_1, \dots, x_N\}$, donde los x_n son independientes e idénticamente distribuidos a partir de una distribución de Bernoulli

$$\text{Bern}(x|\mu) = \mu^x(1 - \mu)^{1-x}.$$

Podemos construir una función de verosimilitud, que es función de μ

$$p(\mathcal{D}|\mu) = \prod_{n=1}^N p(x_n|\mu) = \prod_{n=1}^N \mu^{x_n}(1 - \mu)^{1-x_n} = \mu^S(1 - \mu)^{N-S}$$

con $S = \sum_{n=1}^N x_n$.

Ejemplo: Beta-Bernoulli

Ahora asumamos que μ sigue una distribución de densidad de probabilidad

$$\text{Beta}(\mu, \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \mu^{\alpha-1} (1 - \mu)^{\beta-1}.$$

Aplicando el teorema de Bayes, se obtiene

$$p(\mu|\mathcal{D}) = \frac{p(\mathcal{D}|\mu)p(\mu)}{p(\mathcal{D})}$$

$$p(\mu|\mathcal{D}) \propto p(\mathcal{D}|\mu)p(\mu) = \mu^S (1 - \mu)^{N-S} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \mu^{\alpha-1} (1 - \mu)^{\beta-1}$$

Ejemplo: Beta-Bernoulli

Entonces el posterior queda determinado por

$$\begin{aligned} p(\mu|\mathcal{D}) &\propto \mu^S (1 - \mu)^{N-S} \mu^{\alpha-1} (1 - \mu)^{\beta-1} \\ &\propto \mu^{S+\alpha+1} (1 - \mu)^{N+\beta-S-1} \end{aligned}$$

Podemos decir que

$$p(\mu|\mathcal{D}) = \text{Beta}(S + \alpha, N + \beta - S).$$

Contenido

- 1 Variables aleatorias continuas
 - Funciones de densidad de probabilidad
 - Ejemplos de funciones de densidad de probabilidad
- 2 Vectores aleatorios
 - Preliminares
 - fdp Gaussiana multivariada
- 3 Momentos estadísticos

Función de distribución de vectores aleatorios

- Un **vector aleatorio** es un vector cuyos elementos individuales son variables aleatorias.
- La ley de probabilidad para un vector de variables aleatorias se especifica en términos de función de distribución conjunta

$$F_{X_1, X_2, \dots, X_m}(x_1, x_2, \dots, x_m) = P[(X_1 \leq x_1), (X_2 \leq x_2), \dots, (X_m \leq x_m)].$$

- También se puede especificar en términos de la función de probabilidad de masa (para variables discretas) o en términos de la función de densidad de probabilidad (para variables continuas).
- En lo que sigue se analiza el caso de vectores de variables aleatorias continuas.

Fdp conjunta y marginal

- La fdp conjunta de un vector aleatorio de m dimensiones es la derivada parcial de la función de distribución

$$f_{X_1, X_2, \dots, X_m} = \frac{\partial^m F_{X_1, X_2, \dots, X_m}(x_1, x_2, \dots, x_m)}{\partial x_1 \partial x_2 \dots \partial x_m}$$

- La fdp de una de las variables aleatorias X_i , para $1 \leq i \leq m$ se obtiene integrando con respecto a todas las otras variables aleatorias $X_j \neq X_i$. Por ejemplo, para X_1

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{X_1, \dots, X_m}(x_1, \dots, x_m) dx_2 \dots dx_m$$

- Si se quiere la fdp conjunta de (X_i, X_j) , se marginaliza integrando con respecto a X_k para $k \neq i$ y $k \neq j$. Por ejemplo, la fdp conjunta (X_1, X_2)

$$f_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{X_1, \dots, X_m}(x_1, \dots, x_m) dx_3 \dots dx_m.$$

Fdp condicional

- La fdp condicional de un subconjunto de variables aleatorias del vector aleatorio dado otro subconjunto de variables del mismo vector.
- Ejemplos de fdp condicionales son

$$f_{X_1, X_2, X_3 | X_4}(x_1, x_2, x_3 | x_4) = \frac{f_{X_1, X_2, X_3, X_4}(x_1, x_2, x_3, x_4)}{f_{X_4}(x_4)}$$

$$f_{X_1, X_2 | X_3, X_4}(x_1, x_2 | x_3, x_4) = \frac{f_{X_1, X_2, X_3, X_4}(x_1, x_2, x_3, x_4)}{f_{X_3, X_4}(x_3, x_4)}$$

Valores esperados

- Los valores esperados se calculan usando múltiples integrales.
- Por ejemplo,

$$E\{g(x_1, \dots, x_m)\} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1, \dots, x_m) f_{X_1, \dots, X_m}(x_1, \dots, x_m) dx_1 \dots dx_m.$$

- Igualmente, los valores esperados condicionales se calculan usando funciones de densidad de probabilidad condicionales.

Medias y covarianzas

- Dos momentos estadísticos de importancia en vectores aleatorios son las medias y las covarianzas.
- La media está dada como

$$\mu_{X_i} = E\{X_i\}.$$

- Las covarianzas están dadas como

$$\sigma_{X_i X_j} = E\{X_i X_j\} - \mu_{X_i} \mu_{X_j}.$$

Notación

- La ley de probabilidad para vectores aleatorios puede especificarse de manera concisa usando notación vectorial.
- Supóngase un conjunto de m variables aleatorias X_1, X_2, \dots, X_m .
- Estas variables aleatorias pueden representarse como un vector columna de dimensiones $m \times 1$,

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix}$$

- Un valor específico de \mathbf{X} se denota como $\mathbf{x}^T = (x_1, x_2, \dots, x_m)$.

Fdp con notación vectorial

- Con notación vectorial, la probabilidad conjunta está dada como

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1, \dots, X_m}(x_1, \dots, x_m).$$

- El vector de medias se define como

$$\mu_{\mathbf{X}} = E\{\mathbf{X}\} = \begin{bmatrix} E\{X_1\} \\ E\{X_2\} \\ \vdots \\ E\{X_m\} \end{bmatrix}$$

- La matriz de covarianza está definida como

$$\Sigma_{\mathbf{X}} = E\{\mathbf{X}\mathbf{X}^T\} = \begin{bmatrix} \sigma_{X_1 X_1} & \sigma_{X_1 X_2} & \dots & \sigma_{X_1 X_m} \\ \sigma_{X_2 X_1} & \sigma_{X_2 X_2} & \dots & \sigma_{X_2 X_m} \\ \vdots & \dots & \dots & \vdots \\ \sigma_{X_m X_1} & \sigma_{X_m X_2} & \dots & \sigma_{X_m X_m} \end{bmatrix}$$

Correlación e Independencia

- Se dice que dos componentes del vector aleatorio no están correlacionadas si $\sigma_{X_i X_j} = \sigma_{ij} = 0$.
- Se dice que los componentes del vector aleatorio son independientes si

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^m f_{X_i}(x_i)$$

- Otra notación para la fdp de un vector aleatorio es $p(\mathbf{x})$.

Contenido

1 Variables aleatorias continuas

- Funciones de densidad de probabilidad
- Ejemplos de funciones de densidad de probabilidad

2 Vectores aleatorios

- Preliminares
- fdp Gaussiana multivariada

3 Momentos estadísticos

Fdp Gaussiana multivariada

Se dice que un vector aleatorio \mathbf{X} es Gaussiano multivariado si su fdp sigue la forma

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{m/2} |\Sigma_{\mathbf{X}}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mu_{\mathbf{X}})^{\top} \Sigma_{\mathbf{X}}^{-1} (\mathbf{x} - \mu_{\mathbf{X}}) \right]$$

donde $\mu_{\mathbf{X}}$ es el vector de medias, $\Sigma_{\mathbf{X}}$ es la matriz de covarianza, $|\Sigma_{\mathbf{X}}|$ es el determinante de $\Sigma_{\mathbf{X}}$ y $\Sigma_{\mathbf{X}}^{-1}$ es la matriz inversa de $\Sigma_{\mathbf{X}}$.

Una notación alternativa para la fdp Gaussiana multivariada es la siguiente,

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x} | \mu_{\mathbf{X}}, \Sigma_{\mathbf{X}}) .$$

Fdp Gaussiana multivariada II

Suponga que \mathbf{X} es un vector aleatorio que sigue una fdp Gaussiana. Si se particiona el vector de variables aleatorias \mathbf{X} de la siguiente forma

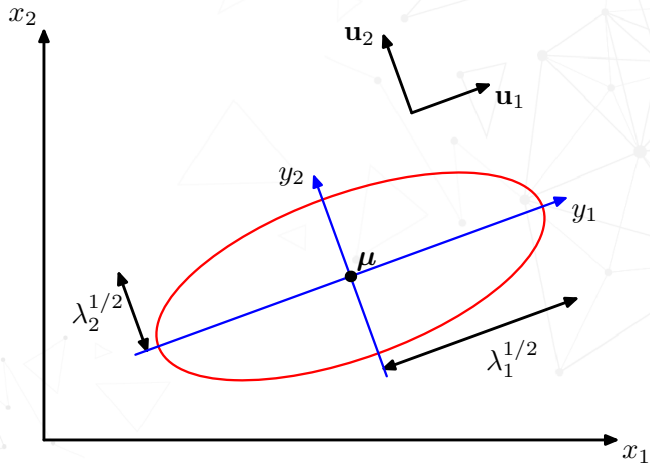
$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \quad \mathbf{X}_1 = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} X_{k+1} \\ X_{k+2} \\ \vdots \\ X_m \end{bmatrix}$$

y

$$\mu_{\mathbf{X}} = \begin{bmatrix} \mu_{\mathbf{X}_1} \\ \mu_{\mathbf{X}_2} \end{bmatrix}, \quad \Sigma_{\mathbf{X}} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

donde $\mu_{\mathbf{X}_i} = E\{\mathbf{X}_i\}$ y $\Sigma_{ij} = E\{\mathbf{X}_i \mathbf{X}_j^T\} - \mu_{\mathbf{X}_i} \mu_{\mathbf{X}_j}^T$, luego \mathbf{X}_1 sigue una fdp Gaussiana multivariada de k dimensiones con media $\mu_{\mathbf{X}_1}$ y la matriz de covarianza Σ_{11} .

Fdp Gaussiana multivariada III



Fdp Gaussiana multivariada IV

- Si $\Sigma_{\mathbf{X}}$ es una matriz diagonal, esto es,

$$\Sigma_{\mathbf{X}} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \sigma_m^2 \end{bmatrix},$$

luego las componentes de \mathbf{X} son independientes (en el caso Gaussiano, la no correlación implica independencia, lo cual no es necesariamente cierto para otras fdps).

- Si $\mathbf{A} \in \mathbb{R}^{k \times m}$ es una matriz de rango k , luego $\mathbf{Y} = \mathbf{A}\mathbf{X}$ sigue una fdp Gaussiana de dimensión k con momentos

$$\mu_{\mathbf{Y}} = \mathbf{A}\mu_{\mathbf{X}},$$

$$\Sigma_{\mathbf{Y}} = \mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}^{\top}.$$

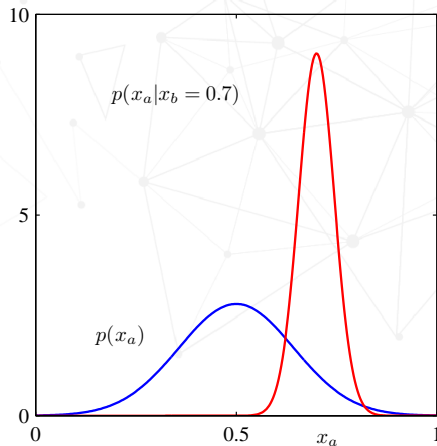
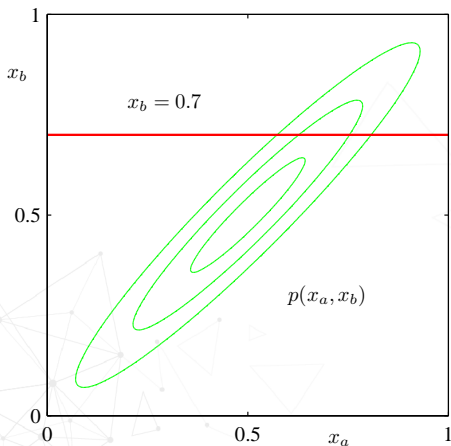
Fdp Gaussiana multivariada V

- Empleando la partición del vector \mathbf{X} como se vió anteriormente, la fdp condicional de \mathbf{X}_1 dado $\mathbf{X}_2 = \mathbf{x}_2$ es una Gaussiana multivariada con los siguientes momentos

$$\begin{aligned}\mu_{\mathbf{X}_1|\mathbf{X}_2} &= E\{\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2\} = \mu_{\mathbf{X}_1} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_{\mathbf{X}_2}), \\ \Sigma_{\mathbf{X}_1|\mathbf{X}_2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.\end{aligned}$$

- Las propiedades anteriores indican que la fdp condicional, la fdp marginal y las transformaciones lineales derivadas de una fdp Gaussiana multivariada conducen a fdp Gaussianas multivariadas.

Ejemplo de dos Variables Discretas



Momentos estadísticos

El n -simo momento de X se representa como $\mathbb{E}[X^n]$, donde,

$$E[X^n] = \begin{cases} \sum_i x_i^n P(X = x_i), & \text{discreta} \\ \int_{-\infty}^{\infty} x^n f(x) dx, & \text{continua} \end{cases}$$

El n -simo momento central de X se representa como $\mathbb{E}[(X - \mu)^n]$, donde,

$$\mu[X^n] = E[(X - \mu)^n] = \begin{cases} \sum_i (x_i - \mu)^n P(X = x_i), & \text{discreta} \\ \int_{-\infty}^{\infty} (x - \mu)^n f(x) dx, & \text{continua} \end{cases}$$

Primer momento estadístico

El primer momento de X se define como

$$\text{Media ó } \mu = E[X] = \begin{cases} \sum_i x_i P(X = x_i), & \text{discreta} \\ \int_{-\infty}^{\infty} x f(x) dx, & \text{continua} \end{cases}$$

Representa la tendencia central de las observaciones o el centro de masas. Es la suma de los productos de la observaciones con sus probabilidades.

Momentos estadísticos: Definiciones

Los siguientes momentos tienen los prefijos “Central” y “Estandarizado”:

- **Central:** momento medido alrededor de la media.
- **Estandarizado:** implica que los momentos son calculados después que las distribuciones son normalizadas.

Segundo momento central

Mide la dispersión de las observaciones alrededor de la media

$$\text{varianza ó } \sigma^2 = E[(X - \mu)^2] = \begin{cases} \sum_i (x_i - \mu)^2 P(X = x_i), & \text{discreta} \\ \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx, & \text{continua} \end{cases}$$

Tercer momento estandarizado (Skewness)

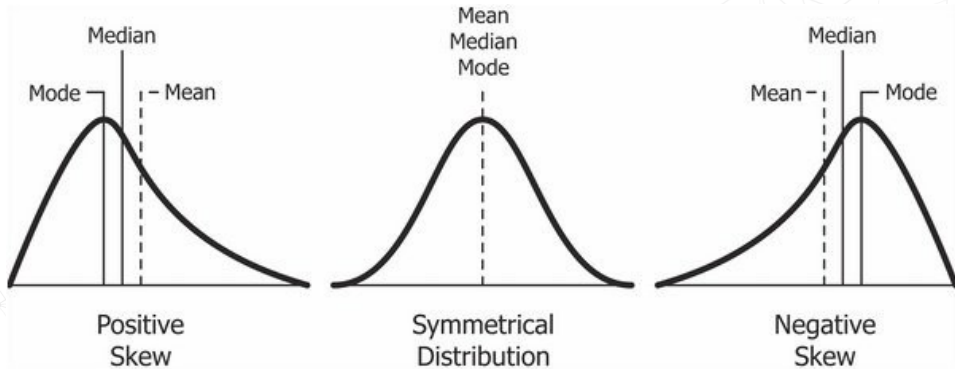
Da una idea de la simetría de la distribución de probabilidades alrededor de la media.

$$\text{Skewness ó } \gamma_1 = E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] = \frac{E[X^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3}$$

Tercer momento estandarizado (Skewness)

- Para tamaños de muestras lo suficientemente grandes (>300), los valores de skewness entre -2 y $+2$ son considerados aceptables con el fin de probar que una distribución es normal univariada.
- Para tamaños muestrales pequeños, se debe convertir las observaciones a valores z y hacer operaciones adicionales (es mejor usar el test K-S o los gráficos Q-Q).

Tercer momento estandarizado (Skewness)



Cuarto momento estandarizado (Kurtosis)

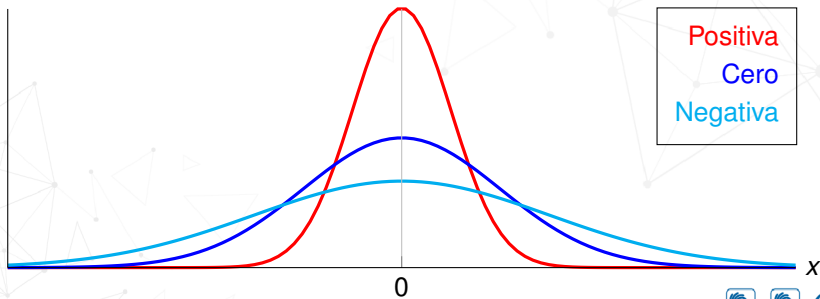
Da una idea de que tan pesada son las colas de una distribución, esto es, que tan frecuente son las desviaciones extremas (outliers) del valor de la media.

$$\text{Kurtosis } \acute{o} \kappa = E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \frac{4^{\text{to}} \text{ momento central}}{\text{Varianza}^2}$$

Para una distribución Gaussiana, $\kappa = \frac{3\sigma^4}{\sigma^4} = 3$

Cuarto momento estandarizado (Kurtosis)

- El cuarto momento se sustrahe 3. Esta kurtosis modificada, se llama exceso de kurtosis.
- El exceso de kurtosis se prefiere sobre la kurtosis por su fácil comparación con la distribución Gaussiana.



Referencias

- Basado en las presentaciones del Prof. Mauricio A. Álvarez, del curso “Procesos Estocásticos”.
- C. Bishop. “Pattern Recognition and Machine Learning”, 2006.
- D.P. Bertsekas et al. “Introduction to Probability”, 2002.
- Medium Article: <https://medium.com/@praveenprashant/the-four-moments-of-a-probability-distribution-6b900a25d0d8>.