Variables Aleatorias y Distribución de Probabilidades

Julio Deride Silva

Área de Matemática
Facultad de Ciencias Químicas y Farmcéuticas
Universidad de Chile

27 de mayo de 2011

Variables Aleatorias y Distribución de Probabilidades

Julio Deride Silva

Área de Matemática
Facultad de Ciencias Químicas y Farmcéuticas
Universidad de Chile

27 de mayo de 2011

Tabla de Contenidos

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Definiciones.

Definición (Variable Aleatoria)

Dado un espacio muestral Ω, se define una variable aleatoria como una función $X : \Omega \to \mathbb{R}$.

Supongamos que la variable aleatoria entrega resultados en un conjunto E. Si se cuenta con una medida de probabilidad $\mathbb{P}$ sobre Ω, se dice que la variable aleatoria X induce una medida de probabilidad $\mathbb{P}_X$ sobre E, $\mathbb{P}_X : \mathcal{P}(E) \to [0, 1]$, dada por $P_X(A) = \mathbb{P}(X^{-1}(A))$, $A \subseteq E$.
También es posible olvidarse de la medida de probabilidad sobre Ω y definir directamente la medida de probabilidad sobre el conjunto de resultados posibles. Esto es lo que se hará en este curso.

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Caso Discreto

Sea X una v.a. discreta con conjunto de resultados posibles:

- $E = \{e_1, e_2, \ldots, e_n\}$ (Caso finito)
- $E = \{e_1, e_2, \ldots, e_n, e_{n+1}, \ldots\}$ (Caso infinito numerable)

Definición (Función de Probabilidad)

Una función de probabilidad para X es una función

$$\begin{array}{rcl} \mathbf{p} : E & \longrightarrow & [0,1] \\ e & \longmapsto & \mathbf{p}(e) = \mathbb{P}(\{e\}) \end{array}$$

que cumple:

$$\sum_{e \in E} \mathbf{p}(e) = 1$$

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Bernoulli(p)

La variable aleatoria discreta X se distribuye según una distribución de Bernnoulli de parámetro $p \in [0, 1]$ (se denota $X \sim Ber(p)$) si X tiene sólo dos resultados posibles, $E = \{0, 1\}$ y el evento $\{1\}$ tiene probabilidad p, es decir:

$$X = \begin{cases} 1 & p \\ 0 & 1 - p \end{cases}$$

Esta variable aleatoria es utilizada para modelar cualquier experimento en donde existan dos resultados posibles y uno se identifique con un éxito (evento $\{1\}$) y otro con fracaso (evento $\{0\}$).

Binomial(n, p)

La variable aleatoria discreta X se distribuye según una distribución de Binomial de parámetros $n \in \mathbb{N}$ y $p \in [0, 1]$ (se denota $X \sim Bi(n, p)$) si X tiene como resultados posibles el conjunto $E = \mathbb{N}$ y el evento $\{k\}$ tiene probabilidad $\mathbf{p}_k$ dada por

$$\mathbf{p}_k = \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

La distribución binomial se utiliza para modelar situaciones en donde interesa contar la cantidad de éxitos, que ocurren con probabilidad p, luego de n intentos.

Poisson(λ)

La variable aleatoria discreta X se distribuye según una distribución de Poisson de parámetro $\lambda > 0$ (se denota $X \sim Poisson(\lambda)$) si X tiene como resultados posibles el conjunto $E = \mathbb{N}$ y el evento $\{k\}$ tiene probabilidad $\mathbf{p}_k$ dada por

$$\mathbf{p}_k = \mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

La distribución de Poisson se utiliza para modelar el número de eventos que ocurren en un lapso fijo de tiempo, si se sabe que estos eventos ocurren en promedio a tasa λ.

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Caso Continuo

Consideremos una variable aleatoria X continua. Supondremos que los resultados posibles estarán en $\mathbb{R}$.

Definición

Una función de densidad de probabilidad es una función tal que

$$\begin{array}{rcl} f : \mathbb{R} & \longrightarrow & [0, \infty) \\ x & \longmapsto & y = f(x) \end{array}$$

que además cumple: $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$

definiciones

Definición

Diremos que una función de densidad de probabilidad f es la densidad de la variable aleatoria X si

$$\mathbb{P}(X \in A) = \int_A f(x)dx \quad A \subseteq \mathbb{R}$$

En particular si A es el intervalo $[a, b]$, se calcula

$$\mathbb{P}(X \in [a, b]) = \int_a^b f(x)dx$$

Densidad de Probabilidad

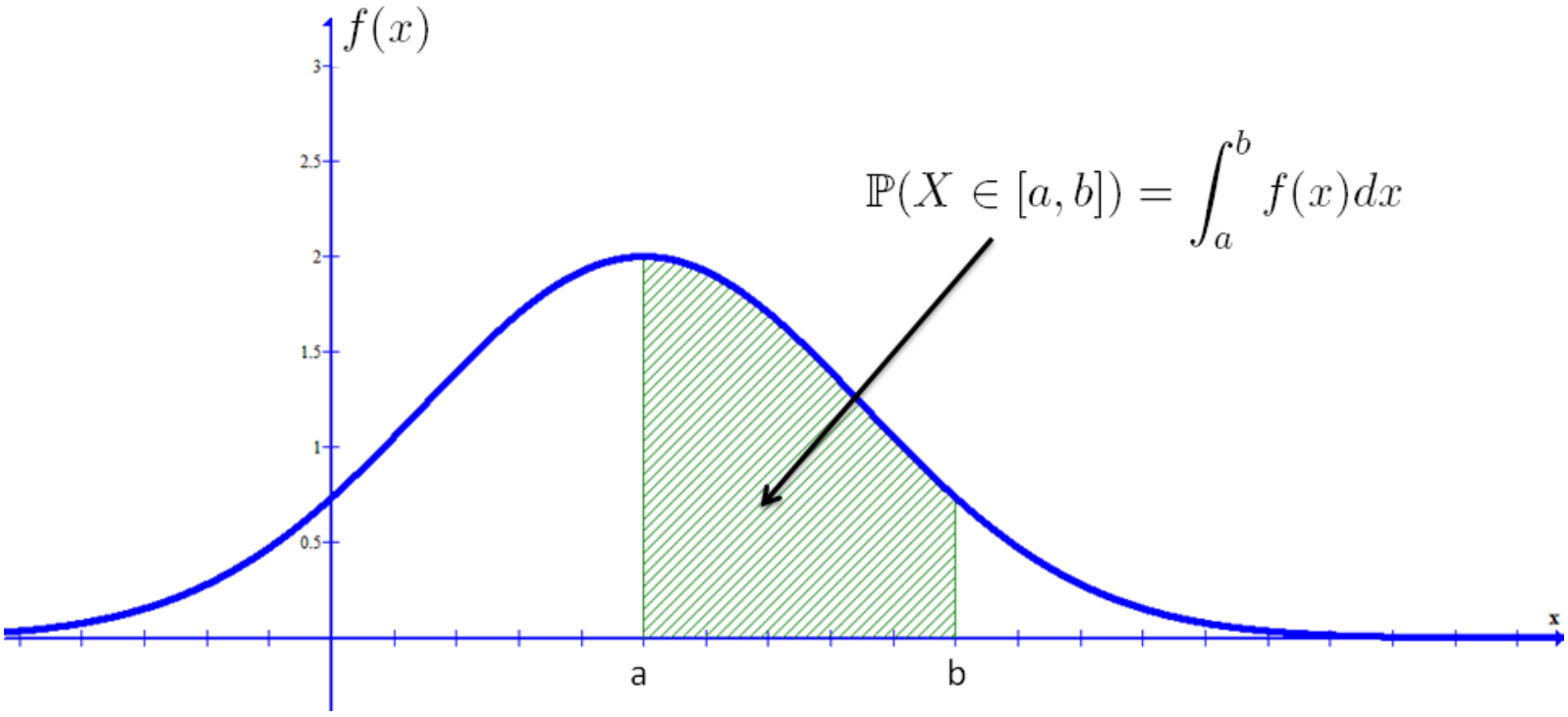

Figura: Función de densidad de Probabilidad

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Distribución Uniforme en $[a, b]$

Una variable aleatoria continua X se dice que sigue una distribución uniforme de parámetros a y b (denotado $X \sim U[a, b]$) si su función de densidad de probabilidad está dada por:

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$$

Distribución Uniforme

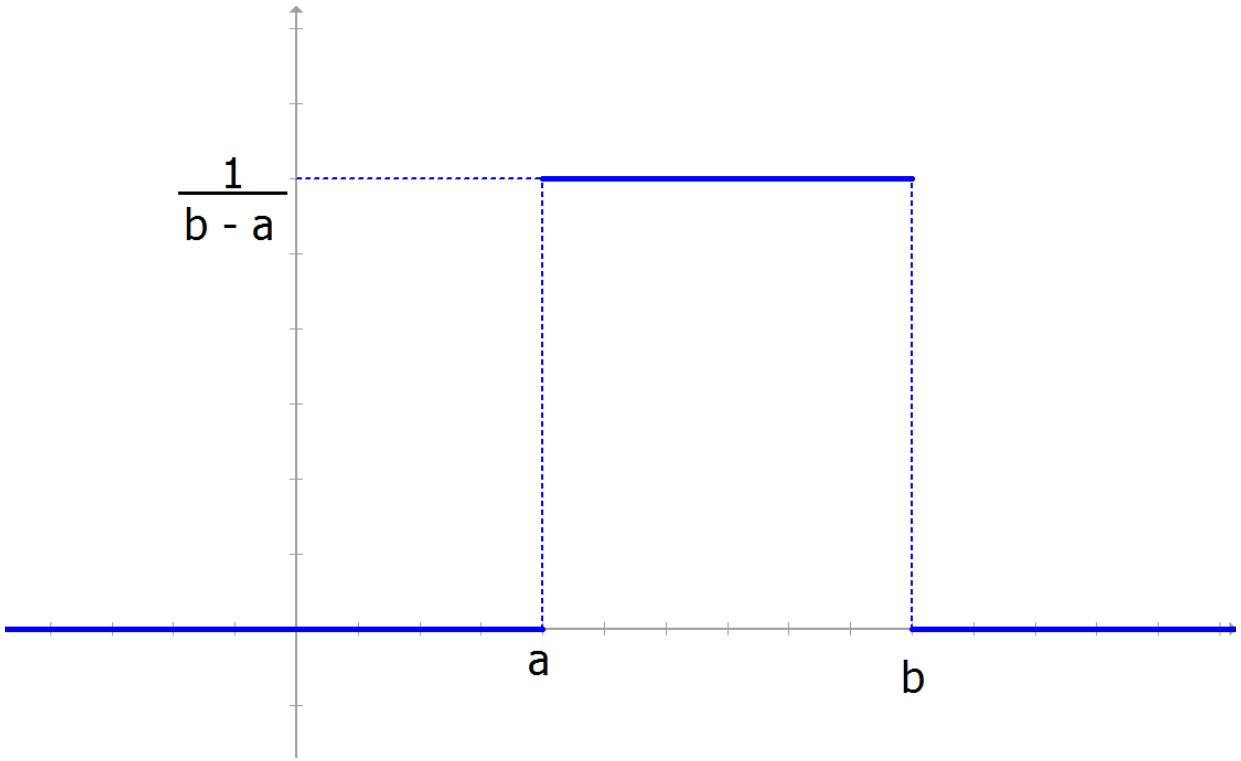

Figura: Función de densidad de una distribución uniforme.

Distribución Exponencial

Una variable aleatoria continua X se dice que sigue una distribución exponencial parámetros λ (denotado $X \sim exp(\lambda)$) si su función de densidad de probabilidad está dada por:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Distribución exponencial

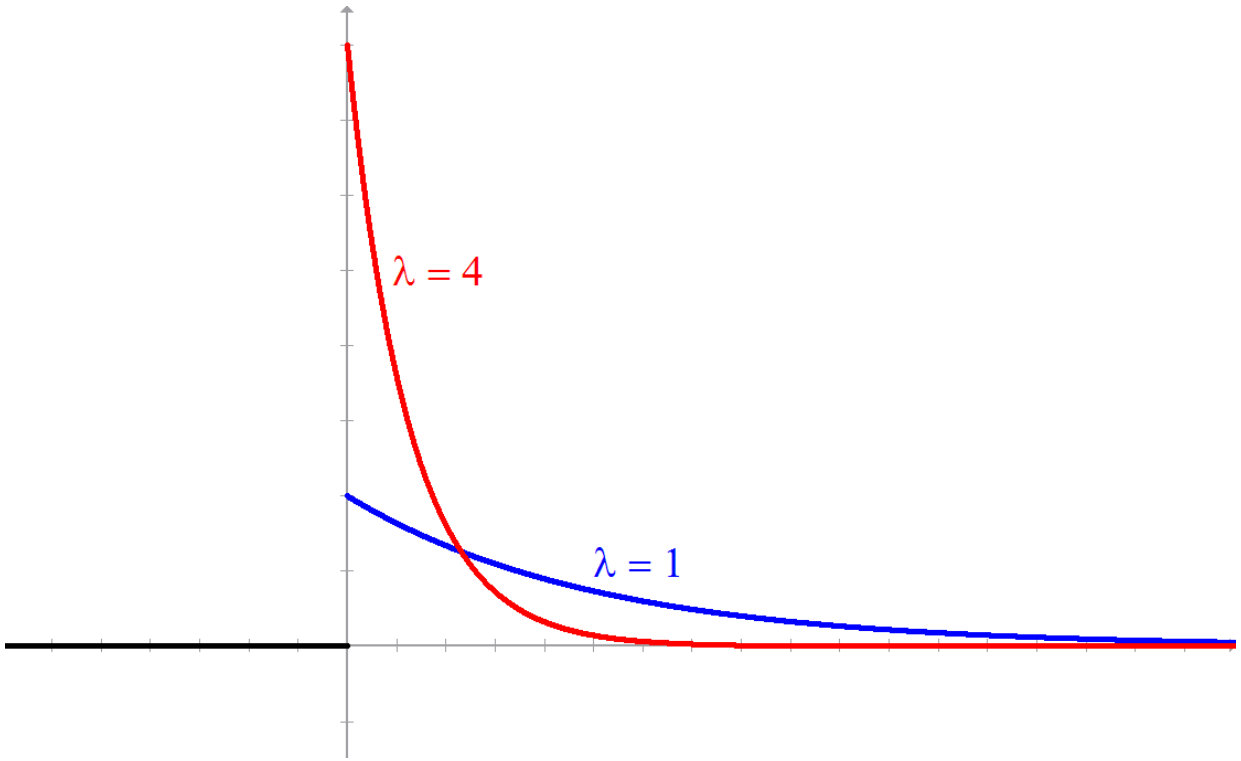

Figura: Función de densidad de una distribución exponencial.

Distribución Normal

Una variable aleatoria continua X se dice que sigue una distribución exponencial parámetros μ y σ^2 (denotado $X \sim \mathcal{N}(\mu, \sigma^2)$) si su función de densidad de probabilidad está dada por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Distribución Normal

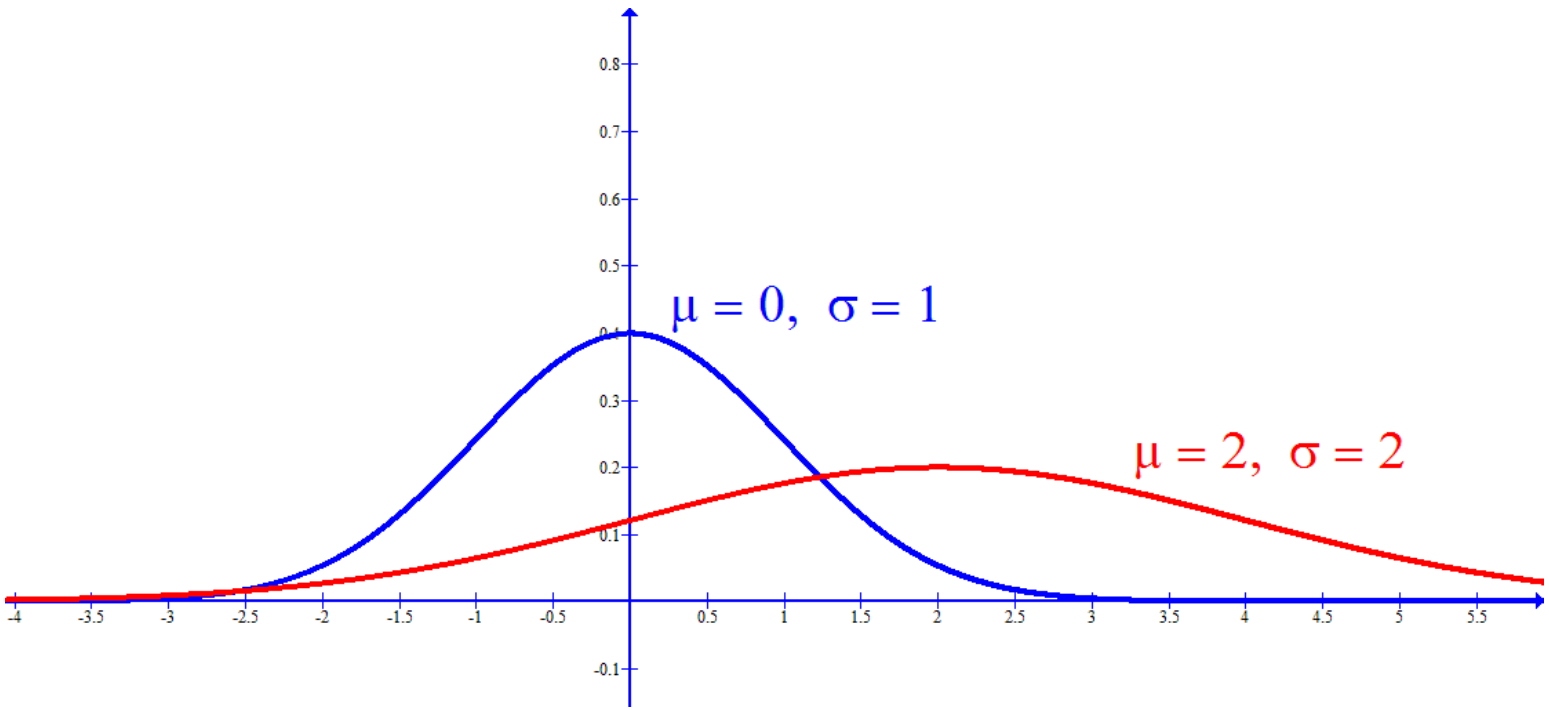

Figura: Densidad de una distribución normal

Distribución Acumulada

Definición (Funcionón de Distribucuión)

Sea X una variable aleatoria continua con función de densidad de probabilidad f. Se define la función de distribución acumulada o simplemente función de distribución como como

$$\begin{aligned} F : \mathbb{R} &\longrightarrow [0,1] \\ a &\longmapsto F(a) = \mathbb{P}((-\infty, a]) = \mathbb{P}(X \leq a) = \int_{-\infty}^{a} f(x)dx \end{aligned}$$

Propiedades

Una función de distribución acumulada posee las siguientes propiedades

1. $\lim_{a \to -\infty} F(a) = 0$
2. $\lim_{a \to \infty} F(a) = 1$
3. F es creciente y continua a la derecha.
4. Por el teorema fundamental del cálculo se cumple que $F'(a) = f(a)$.

Vectores Aleatorios

Definición (Vector Aleatorio)

Una función $X : \Omega \to \mathbb{R}^n$ es un vector aleatorio de n componentes.

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Probabilidad Conjunta

Definición (Probabilidad Conjunta)

Supongamos que el vector aleatorio $X = (X_1, ..., X_n)$ es tal que cada una de sus componentes es una variable aleatoria discreta, con conjunto de resultados posibles E_i, $i = 1, ..., n$. Se define la probabilidad conjunta como una función

$$\begin{aligned} \mathbf{p} : E_1 \times E_2 \times ... \times E_n &\longrightarrow [0,1] \\ (e_1, e_2, ..., e_n) &\longmapsto \mathbf{p}(e_1, e_2, ..., e_n) \end{aligned}$$

además se cumple:

$$\sum_{e_1 \in E_1} \sum_{e_2 \in E_2} \cdots \sum_{e_n \in E_n} \mathbf{p}(e_1, e_2, ..., e_n) = 1$$

Cálculo

Dado un conjunto $A \subseteq \mathbb{R}^n$, se calcula

$$\mathbb{P}(A) = \sum_{(e_1,\ldots,e_n)\in A\cap(E_1\times\ldots\times E_n)} p(e_1,\ldots,e_n).$$

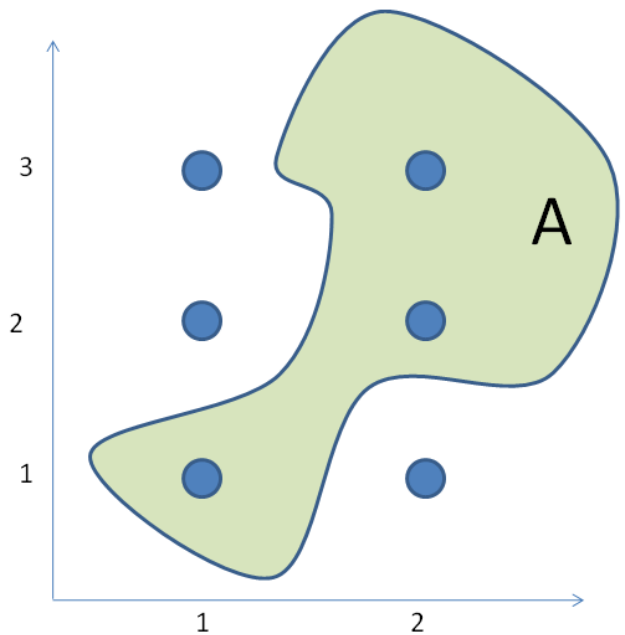

Figura: Probabilidad Conjunta

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Probabilidad Conjunta

Definición (Densidad de Probabilidad Conjunta)

Una función de densidad de probabilidad conjunta es una función tal que

$$\begin{array}{rcl} f : \mathbb{R}^n & \longrightarrow & [0, \infty) \\ (x_1, \ldots, x_n) & \longmapsto & y = f(x_1, \ldots, x_n) \end{array}$$

que además cumple: $\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 dx_2 \ldots dx_n = 1$

Dado un conjunto $A \subseteq \mathbb{R}^n$, se calcula

$$\mathbb{P}(A) = \int_A f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Densidades Marginales

En esta sección se estudia sólo el caso continuo, el caso discreto es similar, reemplazando la función de densidad de probabilidad por la función de probabilidad. Los resultados de esta sección se pueden generalizar para n variables. Por simplicidad se presenta para el caso $n = 2$.

Definición (Densidades Marginales)

Supongamos que un vector (X, Y) tienen densidad conjunta $f(x, y)$. Se definen las densidades marginales por:

- $f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$ *(Marginal de X)*
- $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx$ *(Marginal de Y)*

Independencia

Definición (Independencia)

Dos variables aleatorias X e Y son independientes si

$$\mathbb{P}((X,Y) \in A \times B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B), \qquad \forall A, B \subseteq \mathbb{R}^n$$

Observación: Es posible demostrar que dos variables serán independientes, si y sólo si $f(x,y) = f_X(x) \cdot f_Y(y)$. Es decir, la densidad conjunta es el producto de las densidades marginales.

Outline

1 Variables Aleatorias
- Definición

2 Tipos de Variables Aleatorias
- Caso discreto
- Distribuciones discretas importantes
- Caso continuo
- Distribuciones continuas importantes

3 Vectores Aleatorios
- Caso discreto
- Caso continuo
- Marginales y condicionales
- Esperanza y Varianza

Esperanza

Definición (Esperanza)

Se define la esperanza de una variable aleatoria X como:

$$\mathbb{E}(X) = \begin{cases} \sum_{e \in E} e \cdot \mathbb{P}(e) & X \text{ discreta} \\ \int_{-\infty}^{\infty} x \cdot f(x) dx & X \text{ continuo con densidad } f(x). \end{cases}$$

La esperanza de X representa el centro de masa de la variable aleatoria, si la masa se distribuye según la densidad $f(x)$.

Propiedades:

1. $\mathbb{E}(X) \in \mathbb{R}$, es decir es un número.

Propiedades:

1. $\mathbb{E}(X) \in \mathbb{R}$, es decir es un número.
2. $\mathbb{E}(\cdot)$ es un operador lineal, es decir, dadas X e Y dos V.A. y $\alpha \in \mathbb{R}$, entonces

$$\mathbb{E}(\alpha X + Y) = \alpha \mathbb{E}(X) + \mathbb{E}(Y)$$

Propiedades:

1. $\mathbb{E}(X) \in \mathbb{R}$, es decir es un número.
2. $\mathbb{E}(\cdot)$ es un operador lineal, es decir, dadas X e Y dos V.A. y $\alpha \in \mathbb{R}$, entonces

$$\mathbb{E}(\alpha X + Y) = \alpha \mathbb{E}(X) + \mathbb{E}(Y)$$

3. Sea $g : \mathbb{R} \longrightarrow \mathbb{R}$. Se cumple que
   - $\mathbb{E}(g(X)) = \sum_{e \in E} g(e) \cdot \mathbb{P}(e)$ (Caso discreto)
   - $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(s) \cdot f(s) ds$ (Caso continuo)

Varianza

Definición (Varianza)

Se define la varianza de una variable aleatoria X como

$$\mathbb{V}(X) = \mathbb{E}((X-\mathbb{E}(X))^2) = \begin{cases} \sum_{e \in E} (e - \mathbb{E}(X))^2 \cdot \mathbb{P}(e) & \text{Caso discreto} \\ \int_{-\infty}^{\infty} (s - \mathbb{E}(X))^2 \cdot f(s) ds & \text{Caso continuo co} \end{cases}$$

La varianza mide el grado de dispersión de la función de densidad de probabilidad. Mientras más grande es la varianza, menos concentrada es la función de densidad de probabilidad.

Propiedades:

1. $\mathbb{V}(X) \in \mathbb{R}+$, es decir es un número **positivo**.

Propiedades:

1. $\mathbb{V}(X) \in \mathbb{R}+$, es decir es un número **positivo**.
2. $\mathbb{V}(c) = 0$, para $c \in \mathbb{R}$ constante.

Propiedades:

1. $\mathbb{V}(X) \in \mathbb{R}+$, es decir es un número **positivo**.
2. $\mathbb{V}(c) = 0$, para $c \in \mathbb{R}$ constante.
3. $\mathbb{V}(aX+bY) = a^2\mathbb{V}(X)+b^2\mathbb{V}(Y)+2ab\mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]$.

Covarianza

Definición (Covarianza)

Sean X e Y dos V.A. Se define la covarianza entre X e Y como

$$\mathbf{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))]$$

Se puede probar que esta definición es equivalente a

$$\mathbf{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

La covarianza mide el grado de relación que existe entre las variables X e Y. Un resultado importante es la relación que existe entre la independencia de dos variables aleatorias y la covarianza.

Covarianza e Independencia

Proposición

Si X e Y son independientes, entonces $\mathbf{Cov}(X, Y) = 0$.

Es importante notar el sentido de la implicancia, pues la recíproca no es cierta. Es decir, si $\mathbf{Cov}(X, Y) = 0$ no necesariamente X e Y son independientes[1].

[1] Se cumple cuando X y Y son normales.