Modelos Gráficos Probabilistas

L. Enrique Sucar

INAOE

Sesión 2: Teoría de Probabilidad

"…las reglas mátematicas de la probabilidad no son simplemente reglas para calcular frecuencias de variables aleatorias; son también las únicas reglas consistentes para realizar inferencia de cualquier tipo …"

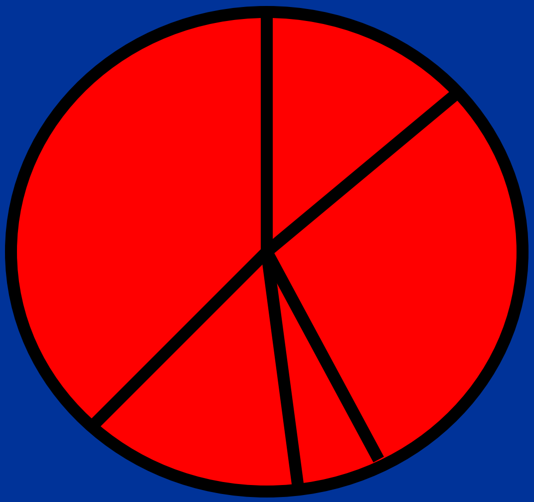
[E. T. Jaynes 2003]

Conceptos de Probabilidad

- Interpretaciones
- Definición y axiomas
- Probabilidad condicional
- Teorema de Bayes
- Independencia e independencia condicional
- Variables aleatorias y distribuciones básicas
- Teoría de información

¿Qué es probabilidad?

- Interpretaciones
- Definición matemática



Interpretaciones

- Clásica – eventos equiprobables
- Lógica – medida de grado de creencia racional (inferencia respecto a evidencia)
- Subjetiva – medida del grado de creencia personal (factor de apuesta)
- Frecuencia – medida del número de ocurrencias con *muchas* repeticiones
- Propensión – medida del número de ocurrencias bajo condiciones repetibles

Interpretaciones

Dos principales enfoques:

- Objetiva (clásica, frecuencia, propensión) – las probabilidades existen y se pueden medir en el mundo real
- Epistemológica (lógica, subjetiva) – las probabilidades tienen que ver con el conocimiento humano, medida de creencia

Justificaciones de Probabilidad

- Argumento del “libro holandés”
- Demostración lógica

Demostración Lógica
(basada en Jaynes)

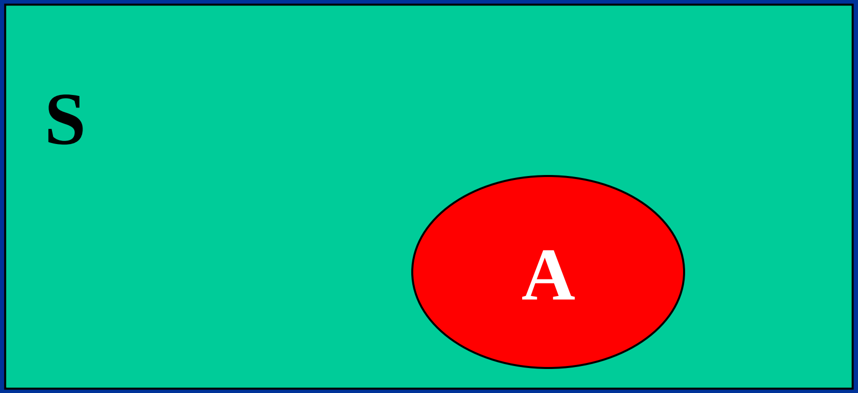
- Condiciones deseables:
  - Los grados de posibilidad son representados por números reales
  - Correspondencia cualitativa con el sentido común
  - Consistencia
    - Si hay varias maneras de llegar a un resultado, todas deben dar el mismo valor
    - Considerar toda la evidencia disponible
    - Representar estados equivalentes de conocimiento con los mismos valores de posibilidad

Reglas básicas

- A partir de las consideraciones anteriores, se derivan las reglas básicas de probabilidad:
  - $0 \leq P(H \mid B) \leq 1$
  - $P(H, D \mid B) = P(H \mid B)\, P(D \mid H, B)$
  - $P(H \mid B) + P(\sim H \mid B) = 1$
- Estas reglas son equivalentes a los axiomas de probabilidad (Kolmogorov)

Definición

- Dado un experimento *E* y el espacio de muestreo *S*, a cada evento *A* le asociamos un número real *P(A)*, el cual es la probabilidad de *A* y satisface los siguientes axiomas



Axiomas

- $0 \leq P(A) \leq 1$
- $P(S) = 1$
- $P(A \cup B \cup C \ldots) = P(A) + P(B) + P(C) + \ldots$

A, B, C … mutuamente exclusivos

Teoremas

- $P(\emptyset) = 0$
- $P(\neg A) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Probabilidad Condicional

$$P(A \mid B) = P(A \cap B) / P(B)$$

- Probabilidad de que ocurra un evento dado que ocurrió otro:
  - Dado que el dado cayó par, cuál es probabilidad de que sea un número primo?
  - Dado que tiene catarro, cuál es la probabilidad de que tenga gripe?

Regla de Bayes

- De la definición de probabilidad condicional se puede deducir:

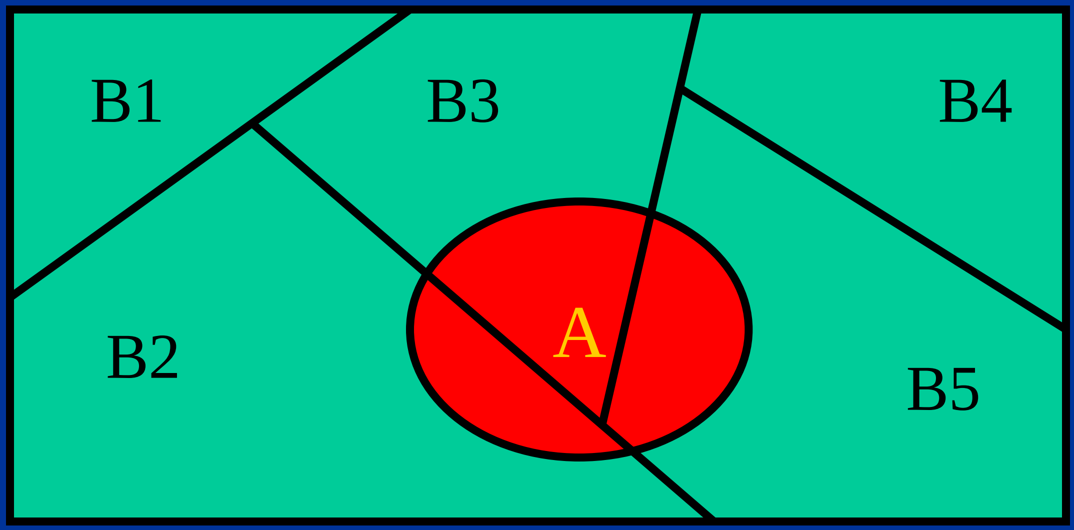
$$P(B \mid A) = P(B)\,P(A \mid B) \,/\, P(A), \text{ dado } P(A) > 0$$

- Esto permite "invertir" las probabilidades, por ejemplo obtener la *P* de una enfermedad dado un síntoma, con conocimiento de la *P* de los síntomas dado que alguien tiene cierta enfermedad

Probabilidad Total

- Dada una partición, *B*, de *S*, la probabilidad de un evento *A* se puede obtener como:

$$P(A) = \Sigma_i\, P(A \mid B_i)\, P(B_i)$$



Teorema de Bayes

- Con la definición de probabilidad total, el teorema de Bayes se puede escribir como:

$$P(B \mid A) = P(B)\, P(A \mid B) \,/\, \Sigma_i\, P(A \mid B_i)\, P(B_i)$$

Eventos independientes

- Dos eventos son independientes si la ocurrencia de uno no altera la probabilidad de ocurrencia del otro:

$$P(A \mid B) = P(A) \text{ ó}$$

$$P(B \mid A) = P(B)$$

- Lo que es equivalente a:

$$P(A \cap B) = P(A)\, P(B)$$

- Independientes $\neq$ mutuamente exclusivos

Independencia condicional

- *A* es condicionalmente independiente de *B* dado *C*, si el conocer *C* hace que *A* y *B* sean independientes:

$$P(A \mid B,C) = P(A \mid C)$$

- Ejemplo:
  - A – regar el jardín
  - B – predicción del clima
  - C – lluvia

Regla de la Cadena

- De la definición de probabilidad condicional, se puede evaluar la probabilidad de $A_1 \cap A_2 \cap A_3 \ldots \cap A_N$ (probabilidad conjunta) como:

$$P(A_1, A_2, \ldots, A_N) =$$

$$P(A_1 \mid A_2, \ldots, A_N)\, P(A_2 \mid A_3, \ldots, A_N) \ldots P(A_N)$$

Variables Aleatorias

- A cada evento A se le asigna un valor numérico $X(A) = k$, de forma que a cada valor le corresponde una probabilidad $P(X = k)$
- X es una variable aleatoria
- Ejemplos:
  – X = Número de águilas en N lanzamientos
  – Y = Número del dado al lanzarlo
  – Z = Número de fallas antes de darle a un blanco

Tipos de Variables Aleatorias

- Discretas: el número de valores de *X* (rango) es finito o contablemente infinito
- Continua: puede asumir todos los posibles valores en cierto intervalo $a - b$, ejemplos:
  - X = temperatura ambiente
  - Y = tiempo en el que falle cierto dispositivo
  - Z = distancia del robot a la pared

Distribución de probabilidad

- Variables discretas: $p(X)$:

  $p(X) \geq 0$

  $\Sigma\, p(X) = 1$

- Variables continuas: $f(x)$:

  $f(x) \geq 0$

  $\int f(x) = 1$

Función acumulativa

- Probabilidad de que la variable X tome un valor menor a x
- Discretas: $P(X) = \Sigma_x\, p(X)$
- Continuas: $F(X) = \int_x f(X)$
- Propiedades:
  - $0 \le F(X) \le 1$
  - $F(X1) \le F(X2)$, si $X1 \le X2$
  - $F(-\infty) = 0$
  - $F(+\infty) = 1$

Estadísticas

- Moda: valor de mayor probabilidad
- Mediana: valor medio (divide el área en 2)
- Promedio: valor "esperado":

  $E(X) = \Sigma_x X\, p(X)$
- Varianza: dispersión

  $\sigma^2(X) = \Sigma_x (X - E(X))^2\, p(X)$
- Desviación estandar

  $\sigma(X) = \sqrt{\sigma^2}$

Variables aleatorias en 2-D

- X y Y son dos funciones que asignan números reales a los eventos en S, entonces (*X*, *Y*) es una variable aleatoria en dos dimensiones
- Propiedades

$p(X,Y) \geq 0$

$\Sigma\,\Sigma\; p(X,Y) = 1$

- Ejemplos:
  - Número de artículos terminados en dos líneas de producción
  - Número de pacientes con cáncer y número que fuma

Probabilidad conjunta, marginal, y condicional

- Probabilidad conjunta:

  $p(X,Y)$

Probabilidad marginal:

$$p(X) = \Sigma_Y \, p(X,Y)$$

- Probabilidad condicional:

$$p(X \mid Y) = p(X,Y) \,/\, p(Y)$$

Independencia y Correlación

- Dos variables aleatorias son independientes si su probabilidad conjunta es el producto de las marginales:

$$p(X,Y) = p(X)\, p(Y)$$

- Correlación: grado de relación lineal entre dos variables aleatorias (diferente a independencia):

$$\rho\ (X,Y) = E\{[(X - E(X)][Y - E(Y)]\} / \sigma_X\, \sigma_{Y,,}$$

$$[-1, 1]$$

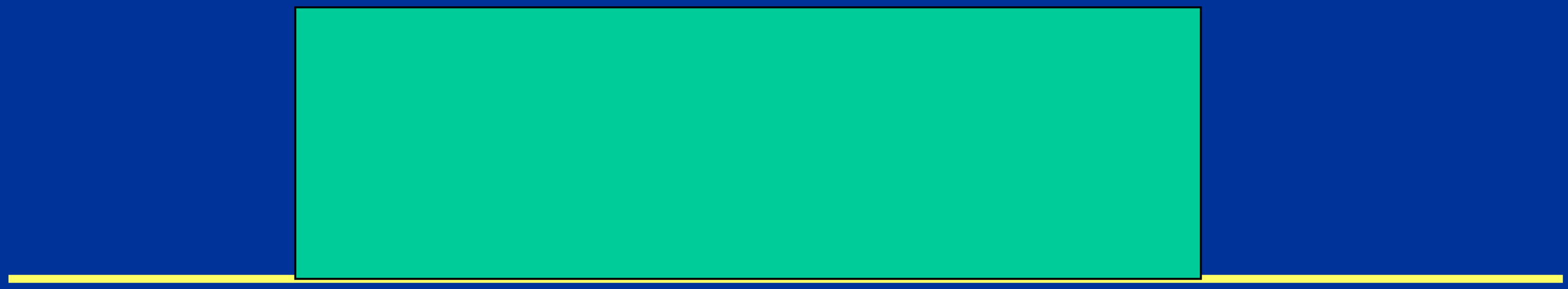
Distribuciones básicas

- Uniforme
- Binomial
- Gaussiana o normal

- Histograma de una variable aleatoria

Uniforme

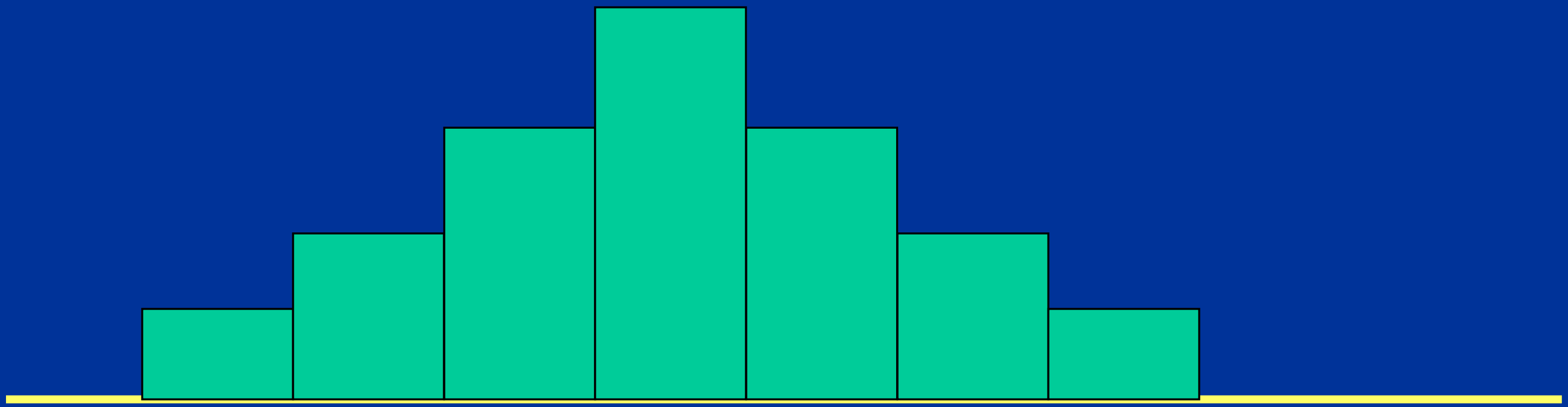
- Todos los valores en el rango son equiprobables



Binomial

- *X* es el número de valores verdaderos en *N* repeticiones de un proceso de Bernoulli con probabilidad *P* de verdadero (éxito)

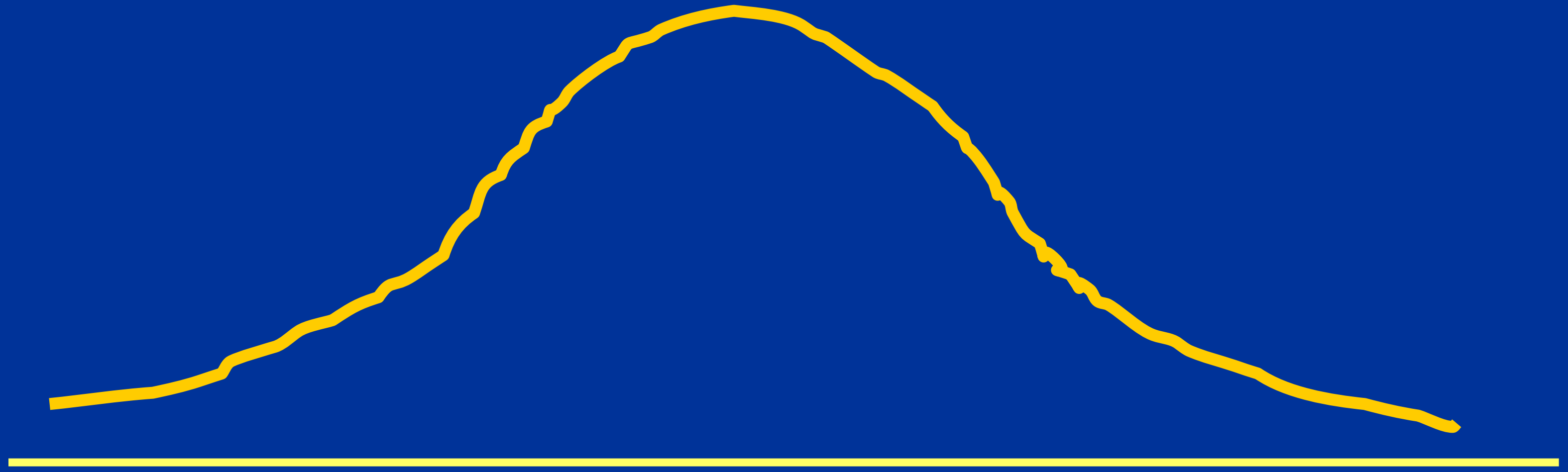
$$P(X=k) = (n\ k)\ p^k\ (1-p)^{n-k}$$



Gaussiana

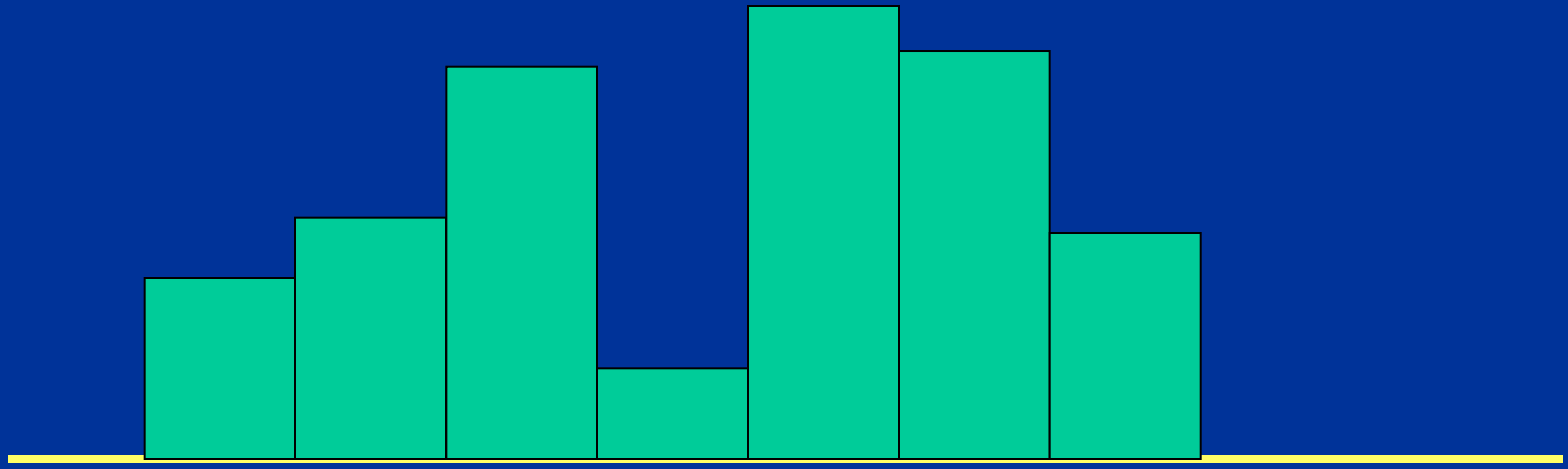
- Aproximación a la binomial con *p*=0.5 y *N* muy grande (corresponde a la suma de muchas variables aleatorias independientes)

$$f(x) = 1/\sigma(2\pi)^{1/2} \exp[-1/2 \, ((x-\mu)/\sigma)^2 \,]$$



Histograma

- Muestra el número de datos por intervalo en forma absoluta o relativa



Referencias

- [Jaynes] Cap. 1, 2
- [Neapolitan] Cap. 2
- [Wesserman] Caps. 1, 2
- Libros básicos de probabilidad, por ej.:
  - Meyer, Introductory Probability and Statistical Applications
  - Waserman, All of statistics

Actividades

- Leer sobre temas de probabilidad (documento sobre interpretaciones en la página)
- Hacer ejercicios de probabilidad en la página del curso (no entregar)
- Ir bajando herramientas de software