Probabilidad conjunta y condicional

Probabilidad condicional	**Probabilidad Conjunta**
p(B/A)= p(AB)/p(A)	p(AB)= p(A)*p(B/A)
p(A/B)=p(AB)/p(B)	p(AB)=p(B)*p(A/B)

Luego:

$\boxed{P(AB)=p(A)*p(B/A)=p(B)*p(A/B)}$ Teorema de Bayes

Recordar que:

si p(B/A)=p(B) son independientes. B no depende de A => p(AB)=p(A)*p(B)
si p(A/B)=p(A) son independientes. A no depende de B =>p(AB)=p(A)*p(B)

Probabilidad Conjunta y Marginales

Probabilidad Conjunta

P(AB)	A_1	A_2	A_i	…	A_n
B_1					
B_2					
B_j			$p(A_iB_j)$		
…					
B_m					

Prob marginal

P(B)
$p(B_j)=\sum_k p(A_kB_j)$

Prob marginal

P(A)			$p(A_i)=\sum_h p(A_iB_h)$		

Propiedades
$\sum_k p(A_k)=1$
$\sum_h p(B_h)=1$
$\sum_{k,h} p(A_kB_h)=1$

Probabilidad Condicional

p(A/B)	A_1	A_2	A_i	…	A_n
B_1					
B_2					
B_j			$p(A_iB_j)/ p(B_j)$		
…					
B_m					

Propiedad
$\sum_k p(A_k/B_j)=1$

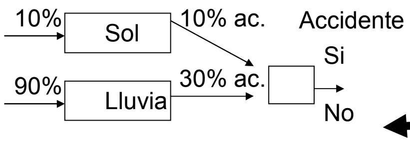
p(B/A)	A_1	A_2	A_i	…	A_n
B_1					
B_2					
B_j			$p(A_iB_j)/ p(A_i)$		
…					
B_m					

$\sum_h p(B_h/A_i)=1$

Importante: Dirección de la flecha para leer la tabla

Ejemplo

Clima=(sol, lluvia)



Probabilidad Condicional Accidente/Clima

P(A/C)	sol	lluvia
Si	**0,10**	**0.30**
No	0.90	0.70

Calcular p(sol/Si)
Es una probabilidad a posteriori "Si se conoce que hubo un accidente cuál es la probabilidad que haya sido con sol o con lluvia.

Teoría de la Información 2011 - Prof. Mariana del Fresno y Prof. Rosana Barbuzza
Facultad de Ciencias Exactas-UNCPBA

Ejemplo

Probabilidad Conjunta p(C A)=p(C)*p(A/C)

P(C A)	sol	lluvia	
Si	p(sol)*p(Si/sol)=0.01	p(lluvia)*p(Si/lluvia)=0.27	p(Si)=0.28
No	p(sol)*p(No/sol)=0.09	p(lluvia)*p(No/lluvia)=0.63	p(No)=0.72
	p(sol)=0.10	p(lluvia)=0.90	

Probabilidad Condicional p(C/A)=p(C A)/p(A)

P(C / A)	sol	lluvia
Si	P(sol y Si)/p(Si)=0.035	p(lluvia y Si)/ p(Si)=**0.965**
No	P(sol y No)/p(No)=0.125	p(lluvia y No)/ p(No)=0.875

Respuesta interpretación

Si hubo accidente es más probable (con 96.5%) que haya sido con lluvia

Ejemplo Muestreo computacional

Ejemplo discreto A={0,1} p(A)={0.1, 0.9} //0 sol 1 lluvia
B={0,1} // 0 accidente 1 no accidente p(B=0/A=0)=0.10; P(B=0/A=1)=0.30
Calcular P(AB) por muestreo computacional

```
float CalcularprobAB()  // A es clima B accidente
{tiradas=0;                       // para el ejemplo
 prob_AB_ant={-1,-1}   prob_AB_act={0,0}
               {-1,-1};                  {0,0}
 exitosAB= {0;0}
              {0,0};
while not converge(prob_AB_ant, prob_AB_act)
{
   A=generaA(); B=generaBdadoA(A)
  exitosAB[A,B]++;
  tiradas++;
  copiar (prob_AB_ant, prob_AB_act);
  for(i=0 to 1)
   for(j=0 to 1)
   prob_AB_act[i,j]=exitosAB[i,j]/tiradas
}
return prob_AB_act;
}
```

```
int generaA()
{probacum={0.1,1} // para este ejemplo
  U=rand ()
  for (i=0 to 1)
     if (U <probacum[ i ])
          return i        //retorna el valor de A {0,1}
}
```

```
int generaBdadoA(int A)
{macumBdadoA={0.1,0.3}
                  { 1  ,  1 } // para este ejemplo
  U=rand ()
  for (i=0 to 1)                  //recorre columna A
     if (U <macumBdadoA[ i, A ])
      return i                     //retorna el valor de B dado A {0,1}
}
```

```
BOOL converge(Mant, Mact)
{ for (i=0 to 1)
    for (j=0 to )
     if (|Mant[i,j] –Mact[i,j]| / Mant[i,j])> ξ
             return false;
 return true;
}
```

Variable Estocástica

- Es un número que se le asocia a cualquier salida de un experimento X(salida)
- Variables aleatorias discretas o continuas

Ejemplos
Se tira un dado, salidas posibles {1,2,3,4,5,6}

X(2)=1, X(4)=1, X(6)=1,
X(1)=2, X(3)=2, X(5)=2

$$X(\text{dado}) \begin{cases} 1 & \text{si dado es par} \\ 2 & \text{en otro caso} \end{cases}$$

Se arroja una moneda tres veces, salidas posibles {CCC,CCS,SCC,CSC,SSC,SCS,CSS,SSS}

Variable estocástica se define como   X(salida)= número de caras
X(CCC)=3,
X(CCS)=X(SCC)=X(CSC)=2
X(SCS)=X(SSC)=X(CSS)=1
X(SSS)=0

Probabilidad

- P(X=x) es la probabilidad que la variable aleatoria X tome el valor x
- Distribución de probabilidad (discreto) donde $\sum_x p(x) = 1$
- Función de densidad de probabilidad (continuo) donde $\int_x f(x)dx = 1$
- La función acumulada F(x) es continua y creciente con máximo 1

$$F(x_2) - F(x_1) = \sum_{x=x_1}^{x_2} p(x)$$

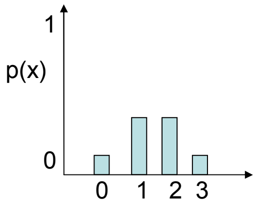
$$F(x_2) - F(x_1) = \int_{x=x_1}^{x_2} f(x)dx$$

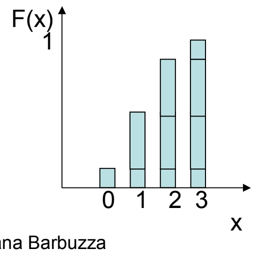
Ejemplo discreto
Se arroja una moneda no pesada tres veces X= número de caras de la salida

X={0,1,2,3} // valores de X   distribución de probabilidad   Función Acumulada

P(X=0)=1/8
P(X=1)=3/8
P(X=2)=3/8
P(X=3)=1/8

$$\sum_{x=0}^{3} p(x) = 1/8 + 3/8 + 3/8 + 1/8 = 1$$

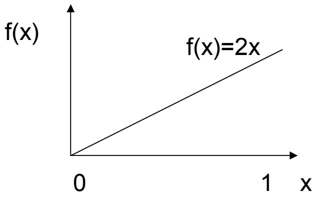


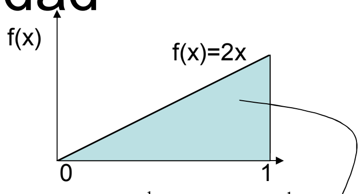


Teoría de la Información 2011 - Prof. Mariana del Fresno y Prof. Rosana Barbuzza
Facultad de Ciencias Exactas-UNCPBA

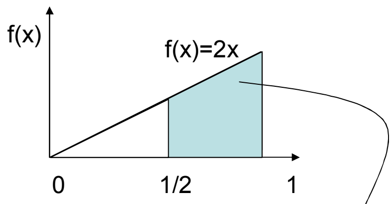
Probabilidad

Ejemplo Continuo





$$F(1)-F(0)=\int_{x=0}^{x=1} 2\,x\,dx = x^2\Big|_0^1 = 1$$



$$F(1)-F(1/2)=\int_{x=1/2}^{x=1} 2\,x\,dx = x^2\Big|_{1/2}^{1} = 1-1/4=3/4$$

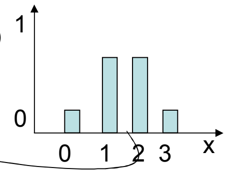
Indicadores

Analítica	Muestreo computacional
• Media $\boxed{<X>=\begin{cases}\sum_x x\,p(x)\\ \int_x x\ f(x)dx\end{cases}}$	$<X>=\dfrac{\sum_i x_i}{N}$ *N* es el número de muestras
• Varianza $\boxed{\sigma_x^2=\sum_x (x-<X>)^2\,p(x)}$ $=\sum_x (x^2-2x<X>+<X>^2)\,p(x)=$ $=\sum_x x^2\,p(x)-2<X>\sum_x x\,p(x)+<X>^2\sum_x p(x)=$ $=<x^2>-2<X>^2+<X>^2$ $\boxed{\sigma_x^2=<x^2>-<X>^2}$	$\sigma_x^2=\dfrac{\sum_i (x_i-\sum_i \frac{x_i}{N})^2}{N}$ *N* es el número de muestras
• Desvío estándar $\sigma_x=\sqrt{\sigma_x^2}=\sqrt{<x^2>-<X>^2}$	$\sigma_x=\sqrt{\dfrac{\sum_i (x_i-\sum_i \frac{x_i}{N})^2}{N}}$

Indicadores (cálculo analítico)

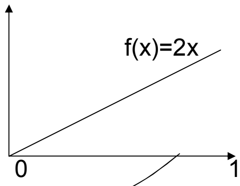
- Promedio o media $< X >= \begin{cases} \sum_x x\, p(x) \\ \int_x x \ f(x) dx \end{cases}$

Ejemplo discreto
X={0,1,2,3} p(x)={1/8,3/8,3/8,1/8}
<X>=0*1/8+1*3/8+2*3/8+3*1/8=12/8=1.5



Ejemplo continuo

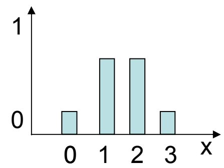
$$< X >= \int_{x=0}^{x=1} x\, 2\, x\, dx = \frac{2}{3} x^3 \Big|_0^1 = 2/3$$



Indicadores (cálculo analítico)

Varianza y Desvío Estándar

Ejemplo discreto
X={0,1,2,3} p(x)={1/8,3/8,3/8,1/8}



Media
<X>=0*1/8+1*3/8+2*3/8+3*1/8=12/8=1.5 caras

Varianza

$$\sigma_x^2 = \sum_x (x - < X >)^2\, p(x) = (0-1.5)^2 * 1/8 + (1-1.5)^2 * 3/8 + (2-1.5)^2 * 3/8 + (3-1.5)^2 * 1/8 = 3/4\ caras^2$$

$$\sigma_x^2 = < x^2 > - < X >^2 = 0^2 * 1/8 + 1^2 * 3/8 + 2^2 * 3/8 + 3^2 * 1/8 - 1.5^2 = 3/4\ caras^2$$

$$\sigma_x = \sqrt{3/4} = 0.86\ caras$$

El desvío estándar mide la dispersión alrededor de la media
A mayor valor de desvío mayor dispersión

Indicadores Cálculo por muestreo computacional

Ejemplo discreto

X={0,1,2,3} p(x)={1/8,3/8,3/8,1/8}. Calcular <X> por muestreo computacional

```
float CalcularMedia()
{suma=0;
 tiradas=0;
 media_ant=-1;
 media_act= 0;

 while not converge(media_ant,media_act)
{
   x=generaX();
  suma=suma+x;
  tiradas++;
  media_ant=media_act;
  media_act=suma/ tiradas
  }
return media_act
}
```

```
int generaX()
{prob={1/8,3/8,3/8,1/8}          // para este ejemplo
  probacum={1/8,4/8,7/8,1}  // para este ejemplo
  U=rand ()
  for (i=0 to 3)
     if (U <probacum[ i ])
          return i        //retorna el valor de x {0,1,2,3}
}
```

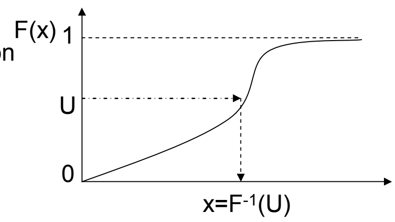
```
BOOL converge(ant, act)
{ if (|ant –act | / ant )< ξ                    // ξ es cte pequeña
                return true;
 return false;
}
```


Generación de muestras

- Generar valores con la computadora de una variable aleatoria X con determinada distribución de probabilidad
- Usamos el método de Inversión

Uso la Función Acumulada

genero aleatorios
$0 \leq U \leq 1$
Calculo $x=F^{-1}(U)$

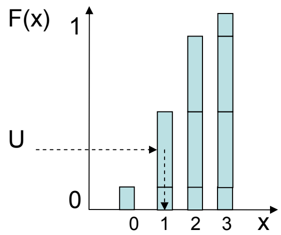


Se obtiene así una muestra de valores independientes de X con la distribución buscada

Ejemplos generación de variable

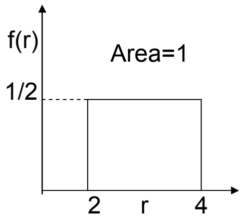
Ejemplo v.a discreta

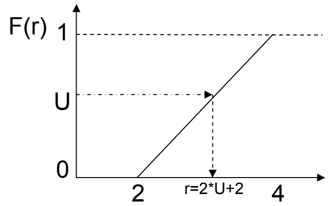
```
int generaX()
{prob={1/8,3/8,3/8,1/8} // para este ejemplo
 probacum={1/8,4/8,7/8,1} // para este ejemplo
 U=rand ()
 for (i=0 to 3)
   if (U <probacum[ i ])
      return i          //retorna el valor de x {0,1,2,3}
}
```



Ejemplo v.a continua

```
int generaR()
{U=rand ()
 r=2*U+2      //para este ejemplo
 return r
// verificar rango de r
//si U=0 retorna 2*0+2=2
// si U=1 retorna 2*1+2=4
}
```



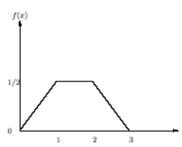


$$F(r)-F(2)=\int_{x=2}^{x=r}\frac{1}{2}\,dx=\frac{1}{2}x\Big|_2^r=\frac{1}{2}r-1\Rightarrow F(r)-F(2)=\frac{1}{2}r-1\Rightarrow F(r)-0=\frac{1}{2}r-1\Rightarrow r=2*F(r)+2$$

U

Teoría de la Información 2011 - Prof. Mariana del Fresno y Prof. Rosana Barbuzza
Facultad de Ciencias Exactas-UNCPBA

Otro ejemplo generación v.a continua



$$f(x)=\begin{cases}\frac{1}{2}x & si \quad 0\le x\le 1\\ \frac{1}{2} & si \quad 1<x<2\\ \frac{1}{2}(3-x) & si \quad 2\le x\le 3\\ 0 & en\ otro\ caso.\end{cases}$$

Caso1 Si F(x)<1/4)

$$\int_0^x\frac{1}{2}x\quad dx=\frac{x^2}{4}\Big|_0^x=\frac{x^2}{4}=u\Rightarrow x=\sqrt{4u}$$

Verificar con f(x):
Si u=0 entonces x=0
Si u=1/4 entonces x=1

Caso2 Si 1/4<F(x)<3/4

$$\frac{1}{4}+\int_1^x\frac{1}{2}\quad dx=\frac{1}{4}+\frac{x}{2}\Big|_1^x=\frac{1}{4}+\frac{x}{2}-\frac{1}{2}=\frac{x}{2}-\frac{1}{4}=u\Rightarrow x=\frac{4u+1}{2}$$

Verificar con f(x)
Si u=1/4 entonces x=1
Si u=3/4 entonces x=2

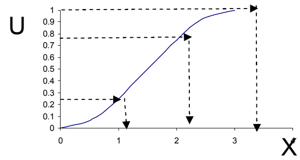
Caso 3 Si 3/4<F(x)<1

$$\frac{1}{4}+\frac{1}{2}+\int_2^x\frac{1}{2}(3-x)dx=\frac{3}{4}+\frac{3x}{2}\Big|_2^x-\frac{x^2}{4}\Big|_2^x=\frac{3}{4}+\frac{3x}{2}-3-\frac{x^2}{4}+1=-\frac{x^2}{4}+\frac{6x}{4}-\frac{5}{4}=u\Rightarrow(x-3)^2-4=-4u\Rightarrow x=3-\sqrt{4-4u}$$

Verificar con f(x)
Si u=3/4 entonces x=2
Si u=1 entonces x=3

int generaX() //para este ejemplo

```
{U=rand ()
If U<1/4 return sqrt(4U)              //según caso 1
else if U<3/4  return (4*U+1)/2       //según caso2
 else return 3-sqrt(4-4U)             //según caso 3
}
```



Teoría de la Información 2011 - Prof. Mariana del Fresno y Prof. Rosana Barbuzza
Facultad de Ciencias Exactas-UNCPBA

Teoría de la Información

Programa del curso

Unidad 1. Tratamiento probabilístico de la información. Variables estocásticas. Distribución de probabilidad. Probabilidad marginal, condicional y conjunta. Teorema de Bayes. Covarianza y correlación. Estimación de variables estocásticas. Muestreo computacional.

Unidad 2. Procesos estocásticos discretos. Cadenas de Markov. Grafos de estados. Equilibrio estadístico. Recurrencia. Fórmulas de Chapman-Kolmogorov. Autocovarianza y autocorrelación.

Unidad 3. Noción de Información. Particiones y árboles. Codificación. Longitud media de código. Concepto de Entropía. Información condicional.

Unidad 4. Fuentes de información y su codificación. Propiedades de las fuentes sin memoria. Fuentes con memoria. Clasificación de códigos. Condición de prefijo y códigos instantáneos. Inecuación de Kraft. Construcción de códigos compactos. Códigos de Shannon, Fano y Huffman. Rendimiento y redundancia de un código. Extensión de fuentes. Propiedad extensiva de la Entropía. Primer Teorema de Shannon.

Unidad 5. Compresión de datos. Métodos de compresión sin pérdida. Esquemas de compresión adaptativos. Codificación dinámica de Huffman. Métodos basados en diccionario. Técnicas de compresión con pérdida. Compresión de texto, imágenes, sonido y video.

Unidad 6. Canales. Transmisión de la información y probabilidades condicionales. Entropías a-priori y a-posteriori. Ruido y pérdida. Clasificación de canales. Canales en serie. Balance de entropías. Información mutua. Capacidad del canal. Probabilidad de error. Segundo Teorema de Shannon.

Unidad 7. Método de máxima entropía. Aplicaciones.

Bibliografía

Abramson N., *Teoría de la Información y Codificación*, Ed. Paraninfo, 1981 Bell T.,
Cleary J., Witten I., *Text Compression*, Prentice Hall, 1990
Chiang C., *An Introduction to Stochastic Processes and their Applications*, R.Krieger Publishing Company, 1968
Cover T., Thomas J., *Elements of Information Theory*, John Wiley & Sons, 1991
Nelson M. *The Data Compression Book*, M&T Books,1992
Papoulis A., *Probability Random Variables and Stochastic Processes*, McGraw-Hill, 1991
Shannon C., Weaver W., *Teoría Matemática de la Comunicación*, Ed.Forja, 1981