FACULTAD
DE INGENIERIA
Universidad de Buenos Aires

Probabilidad y Estadistica B

Resumen Probabilidad y Estadistica B
Segundo cuatrimestre de 2021

Índice

1. Axiomas de Probabilidad

Una probabilidad es una funcion de P que a cada evento A le hace corresponder un numero real P(A) con las siguientes propiedades:

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $A \cap B = \emptyset \Longrightarrow P(A \cap B) = P(A) + P(B)$
4. $P(\bar{A}) = 1 - P(A)$

2. Experimentos con resultados equiprobables

2.1. Laplace

Evento A con M elementos y Ω espacio finito de N elementos:

$$P(A) = \frac{M}{N} = \frac{card(A)}{card(\Omega)}$$

3. Conteo

3.1. Regla del producto

Sirve para conjunto de pares ordenados entre dos conjuntos A y B: (cada uno de A con cada uno de B)

$$A \times B = \{(a, b) : a \in A, b \in B\} = card(A) \cdot card(B)$$

3.2. Permutaciones

Sirve para saber de cuantas formas se pueden ordenar n elementos de un conjunto:

$$n! = 1 \times 2 \times ... \times n$$

3.3. Variaciones

Sirve para subconjuntos ordenados de k elementos, pertenecientes a un conjunto de n elementos, se representa como $(n)_k$:

$$(n)_k = n(n-1)...(n-k+1) = \frac{n!}{(n-k)!}$$

OBS: se hace con el boton nPr

3.4. Combinaciones

Sirve para subconjuntos sin ordenar de k elementos, pertenecientes a un conjunto de n elementos, se representa como $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

OBS: se hace con el boton nCr

3.5. Bolas y urnas

Sirve para bolas indistinguibles y urnas:

$$\#CP = \binom{B + (U-1)}{B}$$

OBS: se hace con el boton nPr

4. Teoremas sobre conjuntos de eventos

4.1. Teorema 1

Sea $A(n)$ una sucesion de eventos tales que $A_n \subset A_{n+1} \forall n$ y $A = \bigcup_{i=0}^{\infty} A_i$:

$$P(A) = \lim_{n \to \infty} P(A_n)$$

4.2. Teorema 2

Sea $A(n)$ una sucesion de eventos tales que $A_{n+1} \subset A_n \forall n$ y $A = \bigcap_{i=0}^{\infty} A_i$:

$$P(A) = \lim_{n \to \infty} P(A_n)$$

4.3. Teorema σ-aditividad

Sea $A = \bigcup_{i=0}^{\infty} A_i \in \mathcal{A}$ con los eventos A_i mutuamente excluyentes 2 a 2, entonces:

$$P(A) = P(\bigcup_{i=0}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

5. Relaciones entre dos eventos: probabilidad condicional e independencia

5.1. Probabilidad condicional

Es la probabilidad que un evento A se de, sabiendo que ya se dio el evento B (A dado B):

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Propiedades de la probabilidad condicional

1. $0 \leq P(A|B) \leq 1, \forall A \in \mathcal{A}$
2. $P(\Omega|B) = 1$
3. Si $A \cap C = \emptyset \to P(A \cup C|B) = P(A|B) + P(C|B)$
4. Si $P(B) > 0$:
   - $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$
   - $P(A \cap B \cap C) = P(A|B \cap C) \cdot P(A|B) \cdot P(C) = P(A|B \cap C) \cdot P(B \cap C)$

Teorema de la probabilidad total

Dada una particion de Ω en $B_1, B_2, ..., B_n$ eventos, dado un evento superpuesto A, la probabilidad de A es:

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

5.2. Independencia de eventos

Dos eventos son independientes cuando:

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

Esto implica que hay la misma proporcion de B en A que en todo Ω y viceversa.

Propiedades de la independencia de eventos

1. Si A y B son independientes, tambien lo son $\bar{A}$ y B, A y $\bar{B}$, $\bar{A}$ y $\bar{B}$
2. $A_1, A_2, ..., A_n$ son independientes sii para cada subconjunto de mas de dos elementos, la interseccion de los sucesos coincide con el producto de las probabilidades.

5.3. Teorema de Bayes

Sean $B_1, B_2, ..., B_k$ una particion de Ω, A un evento de probabilidad positiva:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^{k} P(A|B_i) \cdot P(B_i)}$$

Se deduce de la definicion de probabilidad condicional y el teorema de probabilidad total.

6. Variables aleatorias

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y X: $\Omega \to \mathbb{R}$ una funcion, diremos que X es una variable aleatoria si $X^{-1}(B) \in \mathcal{A}$. Se puede calcular probabilidad como:

$$P(X^{-1}(B)) = P(X \in B)$$

6.1. Funcion de distribucion

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y X una V.A., definimos su funcion de distribucion:

$$F_X(x) = P(X \leq x) \qquad \forall x \in \mathbb{R}$$

Esta funcion se encarga de acumular probabilidad desde $-\infty$ hasta x.
OBS: $P(A < X \leq B) = F_X(B) - F_X(A)$

Propiedades de la funcion de distribucion

1. $F_X \in [0, 1] \quad \forall x \in \mathbb{R}$.
2. F_X es monotona np decreciente.
3. F_X es continua a derecha.
4. $\lim_{x \to -\infty} F_X(x) = 0$ y $\lim_{x \to +\infty} F_X(x) = 1$

6.2. Soporte de una V.A

El soporte de X es:

$$S_X = \{x \in \mathbb{R} : F_X(x) - F_X(x^-) \neq 0 \vee \frac{dF_X(x)}{dx} \neq 0\}$$

6.3. Variables aleatorias discretas

La variable X tiene una distribucion discreta si hay un conjunto $A \in \mathbb{R}$ finito o infinito numerable, tal que $P(X \in A) = 1$.
Sea para cada $x \in A : p_X(x) = P(X = x)$, se verifica que si $B \subset \mathbb{R}$:

$$P(X \in B) = \sum_{x \in B \cap A} p_X(x)$$

Y en particular:

$$\sum_{x \in A} p_X(x) = 1$$

Y la funcion de distribucion es dado un $B = (-\infty, t]$ resulta:

$$P(X \in B) = P(X \leq t) = F_X(t) = \sum_{x \leq t} p_X(x)$$

6.4. Variables aleatorias continuas

Una variable aleatoria es continua si:

1. *a*) El conjunto de valores posibles se compone de todos los numeros que hay en un solo intervalo o una union excluyente de estos.

   b) Ninguno de estos valores tiene un valor de probabilidad positivo $P(x = c) = 0 \quad \forall c \in \mathbb{R}$

2. Se dice que X es una variable continua si existe una funcion $f_X : \mathbb{R} \to \mathbb{R}$, llamada <u>funcion de densidad</u> de probabilidad, que satisface las siguientes condiciones:

   a) $f_X \geq 0 \quad \forall x \in \mathbb{R}$

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

   c) Para cualquier a y b tales que $-\infty < a < b < +\infty$:

   $$P(a < X < b) = \int_a^b f_X(x)dx$$

Teorema

Sea $F_X(x)$ una funcion de distribucion de una V.A.C. (admite derivada), luego:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

OBS: La funcion de densidad solo existe para V.A.C.

6.5. Eventos equivalentes

Dos eventos son equivalentes si acumulan la misma probabilidad. Para V.A.D. significa que ambos eventos tiene la misma probabilidad.

7. Modelos continuos: distribuciones notables

7.1. Distribucion Uniforme

Supongamos una V.A.C. que toma todos los valores sobre un intervalo $[a, b]$. Si $f_X(x)$ esta dada por:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & si \quad a < X < b \\ \\ 0 & e.o.c \end{cases}$$

Se denota como $X \sim \mathcal{U}(a, b)$

7.2. Distribucion exponencial

Una variable aleatoria tiene una distribucion exponencial de parametro $\lambda > 0$ si su funcion de densidad esta dada por:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & si \quad x > 0 \\ \\ 0 & e.o.c \end{cases}$$

Y su funcion de distribucion es:

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & si \quad x < 0 \\ \\ \int_0^\alpha \lambda e^{-\lambda t} dt = 1 - e^{-\lambda \alpha} & e.o.c \end{cases}$$

Propiedades de la exponencial

1. (PERDIDA DE MEMORIA) Si $X \sim \mathcal{E}(\lambda)$ entonces $P(X > t + s | X > t) = P(X > s)$ $\forall \quad t, s \in \mathbb{R}$.
2. Si X es una V.A.C. y $P(X > t + s | X > t) = P(X > s)$ $\forall \quad t, s \in \mathbb{R}^+$ entonces existe $\lambda > 0$ tal que $X \sim \mathcal{E}(\lambda)$

7.3. Funcion de Riesgo (para V.A.C.)

Dada la funcion intensidad de fallas $\lambda(t)$, la funcion de distribucion es:

$$F(t) = 1 - e^{-\int_0^\infty \lambda(s) ds} \qquad \text{si } t > 0$$

7.4. Distribucion Gamma

Se dice una V.A tiene distribucion Gamma de parametros λ y k si su funcion de densidad es:

$$f_X(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} \qquad \text{si } \{x > 0\}$$

7.5. Distribucion normal estandar

La V.A. X que toma los valores $-\infty < x < +\infty$ tiene una distribucion normal estandar si su funcion de densidad es:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

Para calcular probabilidades de esta distribucion hay que mirar la tabla o integrar numericamente.

7.6. Cuantil de una V.A

Un cuantil α de X es cualquier numero x_α tal que :

$$P(X < x_\alpha) \leq \alpha \text{ y } P(X > x_\alpha) \leq 1 - \alpha$$

8. Funciones de variables aleatorias

Sea $Y = g(X)$ con X una variable aleatoria:
Si X es una V.A.D., Y sera discreta con:

$$p_Y(y) = P(Y = y) = \sum_{x \in B} p_x(x) \qquad \text{con} \quad B = \{x \in \mathbb{R} : g(x) = y\}$$

Y en general:

$$F_Y(y) = P(Y \leq y) = P(g(x) \leq y)$$

Y con esta ultima se calcula la probabilidad $\forall y \in \mathbb{R}$

8.1. Simulacion

Sabiendo la distribucion de X y teniendo una variable aleatoria U para generar valores al azar, sabiendo su distribucion, entonces se busca una $F_U(u_i) = F_X(x_i)$, de donde se puede despejar x_i en funcion de u_i. Este despeje se puede hacer mediante la INVERSA GENERALIZADA:

$$F_X^{-1}(u) = min\{x \in \mathbb{R} : F_X(x) \leq u\} \qquad \text{con } u \in (0,1)$$

Teorema

Si f es una funcion que cumple:

- Ser no decreciente.
- $\lim_{x \to +\infty} F(x) = 1$ y $\lim_{x \to -\infty} F(x) = 0$.
- Continua a derecha.

$\Rightarrow X = F^{-1}(U)$ con $U \sim \mathcal{U}(0,1)$, se tiene que X es una V.A. cuya funcion de de distribucion es la funcion F dada.

9. Truncamiento

Sea X una V.A con $F_X(x) = P(X \leq x)$

$$F_{X|X \in A}(x) = P(X \leq x | X \in A) \qquad = \frac{P(X \leq x, X \in A)}{P(X \in A)}$$

Si X es continua, $f_X(x) = \frac{dF_X(x)}{dx}$

$$\Rightarrow f_{X|X \in A}(x) = \frac{d}{dx} F_{X|X \in A(x)} = \frac{f_X(x) \mathbb{1}\{X \in A\}}{P(X \in A)}$$

10. Vectores Aleatorios

$\mathbb{X} = (X_1, X_2, X_3, ..., X_n)$ es un vector aleatorio de dimension n si para cada $j = 1, ..., n$; $X_j : \Omega \to \mathbb{R}$ es una V.A.

10.1. Funcion de distribucion de un vector aleatorio

Sea $\mathbb{X}$ un vector aleatorio de dimension n, definimos la funcion de distribucion de $\mathbb{X}$ como:

$$F_{\mathbb{X}}(\bar{x}) = P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3, ..., X_n \leq x_n)$$

10.2. Propiedades del vector aleatorio ($\mathbb{X} = (X, Y)$)

1. $\lim_{x,y \to \infty} F_{\mathbb{X}}(x) = 1,$
2. $F_{\mathbb{X}(x)}$ es monotona no decreciente en cada variable.
3. $F_{\mathbb{X}(x)}$ es continua a derecha en cada variable.
4. $P((X, Y) \in (a_1, b_1)x(a_2, b_2)) = F_{\mathbb{X}(b_1, b2)} - F_{\mathbb{X}(b_1, a2)} - F_{\mathbb{X}(a_1, b2)} + F_{\mathbb{X}(a_1, a2)}$

10.3. Funcion de probabilidad de vectores aleatorios discretos (probabilidad conjunta)

Sean X e Y dos V.A.D definidas en el espacio muestral Ω de un experimento. La funcion de probabilidad conjunta se define para cada par de numeros (x, y) como:

$$p_{X,Y}(x, y) = P(X = x, Y = y)$$

Debe cumplirse que:

1. $p_{X,Y}(x, y) \geq 0$
2. $\sum_x \sum_y p_{X,Y}(x, y) = 1$

10.3.1. Funciones de probabilidad marginales y de conjuntos

Para el caso de las variables aleatorias recien mencionadas, sus funciones de probabilidad marginales estan dadas por:

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

Para el caso de cualquier conjunto A compuesto por pares de valores (x, y) entonces:

$$P((X, Y) \in A) = \sum_{(x,y) \in A} \sum p_{X,Y}(x, y)$$

10.4. Funcion de densidad de un vector aleatorio continuo

Sean X e Y V.A.C una funcion de denisidad de probabilidad conjunta $f_{X,Y}(x, y)$ de estas dos variables es una funcion que satisface:

1. $f_{X,Y}(x, y) \geq 0$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$

10.4.1. Funciones de densidad marginales

Para calcular las funciones de densidad marginales de X e Y:

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

10.5. Independecia de vectores aleatorios

Sea (X, Y) un vector aleatorio, las variables aleatorias X e Y son independientes si y solo si:

$$P((X \in A) \cap (Y \in B)) = P(X \in A) \cdot P(Y \in B) \quad \forall A, B$$

10.5.1. Propiedades de la independencia de vectores aleatorios

1. Se dice que $X_1 ... X_n$ son V.A independientes sii

$$F_{X_1 ... X_n}(x_1, ..., x_n) = F_{X_1}(x_1) \cdot ... \cdot F_{X_n(x_n)}$$

2. Se dice que las V.A discretas $X_1, ..., X_n$ independientes sii

$$p_{X_1,...,X_n}(x_1, ..., x_n) = p_{X_1}(x_1) \cdot ... \cdot p_{X_n}(x_n)$$

3. Se dice que las V.A continuas $X_1, ..., X_n$ son independientes sii

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_{X_1}(x_1) \cdot ... \cdot f_{X_n}(x_n)$$

11. Momentos

11.1. Esperanza

Es el promedio ponderado de los valores que puede tomar una V.A. (çentro de masa") Sea X una V.A.D con funcion de probabilidad $p_X(x)$, el valor esperado (o media) de X es:

1. Para discretas:

$$E(X) = \sum_{x \in R_x} x \cdot p_X(x)$$

2. Para continuas:

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

11.1.1. Propiedades

1. El valor de la esperanza de cualquier funcion $h(x)$ (una V.A.) se calcula como:

   a) Para discretas:

$$E(h(x)) = \sum_{x \in R_x} h(x) \cdot p_x(x)$$

   b) Para continuas:

$$E(h(x)) = \int_{-\infty}^{\infty} h(x) \cdot f_x(x)$$

2. Sea X una V.A. con E(X) $= \mu$ si $h(x) = aX + b \to E(h(X)) = a\mu + b$

11.1.2. CASO GENERAL

Sea X una V.A. con funcion de distribucion $F_X(x) = P(X \geq x)$ si $h(X)$ es una funcion cualquiera de X, si definimos A como el conjunto de atomos (valores de X que concentren masa positiva), entonces:

$$E[h(X)] = \sum_{x \in A} h(x) \cdot P(X = x) + \int_{\mathbb{R}/A} h(x) \cdot F'_X(x) dx$$

11.2. Esperanza condicional

$$E[X|X \in A] = \frac{E[X\mathbb{1}\{X \in A\}]}{P(X \in A)}$$

Si despejo, y pienso en una particion tenemos:

$$E(X) = E[X|X \in A] \cdot P(X \in A) + E[X|X \in \bar{A}] \cdot P(X \in \bar{A})$$

11.2.1. Propiedad

Otra manera para calcular la esperanza que puede ser util:

$$E(X) = \int_0^\infty (1 - F_X(x))dx - \int_{-\infty}^0 F_X(x)dx$$

11.3. Varianza

Sea X una V.A, definimos la varianza X como:

$$Var(X) = E[(X - E(X))^2]$$

11.3.1. Propiedad

$$Var(X) = E[(X - E(X))^2] = E[X^2 - 2XE(X) + E(X)]$$

Si $E(X) = \mu$

$$\rightarrow Var(X) = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - \mu^2 = E(X^2) - E(X)^2$$

11.3.2. Desvio estandar

Se define como la raiz cuadrada de la varianza de una V.A.:

$$\sigma_x = \sqrt{Var(X)}$$

11.3.3. Mediana

Es el valor de X que acumula una probabilidad de 0.5 (es el cuantil 0.5) : $X/F_X(x) = 0{,}5$

11.3.4. Moda

Es el valor de X con mayor probabilidad.

11.4. Esperanza de una funcion de un vector aleatorio

La esperanza de una funcion $h(X,Y)$ esta dada por:

1. Si (X,Y) es un vector aleatorio discreto:

$$E(h(X,Y)) = \sum_x \sum_y h(x,y) p_{X,Y}(x,y)$$

2. Si (X,Y) es un vector aleatorio continuo:

$$E(h(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) f_{X,Y}(x,y) dxdy$$

11.4.1. Propiedades de Orden

Sea $X = (X_1, ..., X_n)$ un vector aleatorio, $g : \mathbb{R}^k \to \mathbb{R}$ una funcion, tenemos que:

1. Si $X > 0 \to E(X) > 0$
2. Si $g(x) > 0 \to E(g(X)) > 0$
3. Sea $h(x) > g(x) \to E(h(X)) > E(g(X))$
4. $E(|X|) \geq E(X)$
5. $E(|XY|) \leq \sqrt{E(X^2)E(Y^2)}$

11.4.2. Propiedades importantes

1. $$E[\sum_{i=1}^{n} a_i X_i] = \sum_{i=1}^{n} a_i E(X_i)$$
2. Si $X_1, ..., X_n$ son independientes entonces:
$$E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i)$$

11.5. Covarinza

Sean X e Y dos V.A.:

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

11.5.1. Propiedades de la Covarinza

1. $Cov(X,Y) = E(X \cdot Y) - E(X)E(Y)$
2. Si X e Y son independientes entonces $E(X \cdot Y) = E(X)E(Y) \to Cov(X,Y) = 0$
3. $Cov(a + bX, c + dY) = b \cdot d \cdot Cov(X,Y)$
4. $Cov(X + Y, Z) = Cov(X,Z) + Cov(Y,Z)$
5. $Var(X + Y) = Var(x) + Var(Y) + 2Cov(X,Y)$

11.5.2. Coeficiente de correlacion

Entre las V.A X e Y esta dado por:

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

$\to -1 \leq \rho_{XY} \leq 1$
Propiedad:

$$|\rho_{XY}| = 1 \leftrightarrow P(aX + b = Y) = 1$$

12. Prediccion

Sea Y una V.A., $\mathbb{X} = (X_1, ..., X_n)$ un vector aleatorio, existira alguna funcion $g(\mathbb{X})$ que nos sirva para predecir a Y. Para encontrar dicha funcion se calcula el error cuadratico medio:

$$ECM = E[(Y - g(\mathbb{X}))^2]$$

12.1. Los mejores predictores

1. Constante: $E(X)$
2. Lineal: Recta de regresion de Y basada en X

$$g(X) = \hat{Y} = \frac{Cov(X,Y)}{Var(X)}(X - E(X)) + E(Y)$$

13. Desigualdades

13.1. Desigualdad de Markov

Sea $h : \mathbb{R} \to \mathbb{R}^+$ tal que h es par, y restringida a R^+ es creciente, y sea X una V.A, tal que $E(h(X))$ existe, entonces $\forall t \in \mathbb{R}$,

$$P(|X| \geq t) \leq \frac{E[h(X)]}{h(t)}$$

Si ademas X es no negativa, $\forall a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a}$$

13.2. Desigualdad de Tchevychev

Sea X una V.A. con varianza finita, $\forall k > 0$:

$$P(|X - E(X)| \geq k) \leq \frac{Var(X)}{k^2}$$

(desigualdad de Markov si $Y = X - E(X)$ y $h(t) = t^2$)

14. Funcion de variable aleatoria (cambio de variable)

Sea X una variable aleatoria, sea $Y = g(x)$ una funcion de la variable X, buscamos encontrar la distribucion de la variable Y.

14.1. Metodo de eventos equivalentes

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

14.2. Metodo del Jacobiano (para vectores aleatorios)

Suponga que Y_1 e Y_2 con V.A. continuas con funcion de densidad conjunta $f_{Y_1,Y_2}(y_1, y_2)$ y que para todo (y_1, y_2) tal que $f_{Y_1,Y_2}(y_1, y_2) > 0$, $u_1 = h_1(y_1, y_2)$, $u_2 = h_2(y_1, y_2)$ es una transformacion 1 a 1 de (y_1, y_2) y (u_1, u_2) con inversa $y_1 = h_1^{-1}(u_1, u_2)$, $y_2 = h_2^{-1}(u_1, u_2)$. Si las inversas tienen derivadas parciales continuas respecto de u_1 y u_2 con jacobiano J, entonces:

$$f_{U_1,U_2}(u_1, u_2) = \left.\frac{f_{Y_1,Y_2}(y_1, y_2)}{|J|}\right|_{h_1^{-1}, h_2^{-1}}$$

15. Variables aleatorias condicionadas

Sea X e Y variables aleatorias discretas con $p_X(x) > 0$, la funcion de probabilidad de Y dada que $X = x$ es:

$$p_{Y|X=x}(y) = P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)}$$
$$\rightarrow p_{Y|X=x}(y) = \frac{p_{X,Y}(x, y)}{p_X(x)}$$

Se define como $p_{Y|X=x}(y) = 0$ cuando $p_X(x) = 0$
Recordar:

$$p_{X,Y}(x, y) = p_{Y|X=x}(y)p_X(x)$$
$$p_{X,Y}(x, y) = p_{X|Y=y}(x)p_Y(y)$$

Sea (X, Y) un vector aleatorio continuo con densidad conjunta $f_{X,Y}(x, y)$ y densidad marginal $f_X(x)$, entonces para cualquier valor de X con el cual $f_X(x)$, la funcion de densidad de la variable condicionada Y dada $X = x$ es:

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

Si $f_X(x) = 0$ entonces se define como $f_{Y|X=x} = 0$

15.1. Propiedades

1. Sea X e Y V.A. discretas tales que $p_{Y|X=x}(y) = p_y(y) \quad \forall x \in \mathbb{R}$
   $\rightarrow X$ e Y son independientes.
2. $f_{X,Y}(x, y) = f_{Y|X=x}(y) \cdot f_X(x)$

16. Mezcla de variables aleatorias: Bayes para la mezcla

Sea M una V.A.D. con soporte finito y X una V.A.C., de manera que conozco la funcion de probabilidad de M y la funcion de densidad de las variables aleatorias $X|M = m, \quad \forall m \in R_M$. La funcion de probabilidad de M dado que $X = x$ sera:

$$p_{M|X=x}(m) = \frac{f_{X|M=m}(x) \cdot P(M = m)}{\sum_{i \in R_M} f_{X|M=i}(x) \cdot P(M = i)}$$

17. Esperanza condicional

Si $Y|X = x$ es una V.A.D.

$$\rightarrow E[Y|X = x] = \sum_{y \in R_{Y|X=x}} y \cdot p_{Y|X=x}(y)$$

Si $Y|X = x$ es una V.A.C.

$$\rightarrow E[Y|X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y|X=x}(y)$$

Ambas son funcion de x y se las llama Funcion de Regresion ($\phi(x)$)
De aca desprende la definicion de Esperanza Condicional:
Si llamamos $\phi(x) = E[Y|X = x]$ a la esperanza de la variable condicionada Y dado que $X = x$, luego $\phi : Sop(X) \rightarrow \mathbb{R}$

Vamos a definir una variable aleatoria llamada Esperanza Condicional de Y dada X, denotada por $E[Y|X]$, como $\phi(X) = E[Y|X]$ (cuidado!!!! es una V.A, no una esperanza)
Otra definicion:
La V.A. Esperanza Condicional de Y dada X se define como $\phi = E[Y|X]$, con ϕ una funcion medible tal que $E[(Y - \phi(x)) \cdot t(X)] = 0$ para toda funcion t medible $t : Sop(x) \to \mathbb{R}$ tal que $Y \cdot t(X)$ tiene esperanza finita. Definimos $\phi(x) = E[Y|X = x]$

17.1. Propiedades

1. $E[E[Y|X]] = E[Y]$
2. Sea X e Y variables aleatorias, s y r funciones medibles tales que las variables aleatorias $r(X) \cdot s(Y)$; $r(X)$ y $s(Y)$ tiene esperanza finita, entonces

$$E[r(X) \cdot s(Y)|X] = r(X)E[s(Y)|X]$$

3. Sean Y_1 e Y_2 V.A con esperanza finita

$$E[aY_1 + bY_2|X] = aE[Y_1|X] + bE[Y_2|X]$$

4. $E[Y|X] = E[Y]$ si X e Y son independientes
5. $E[r(X)|X] = r(X)$

OBS: La esperanza condicional es de Y dado X es la funcion de la V.A. X que mejor predice a Y.

18. Varianza Condicional

Sean X e Y variables aleatorias con $Var(Y)$ finita, la varianza de $Y|X = x$ sera:

$$Var(Y|X = x) = E[(Y - E(Y|X = x))^2|X = x]$$

Llamaremos $T(x) = Var(Y|X = x)$, $T : Sop(X) \to \mathbb{R}$ llamaremos varianza condicional de Y dada X a la V.A:

$$T(X) = Var(Y|X) = E[(Y - E(Y|X))^2|X]$$

Desarrollando el cuadrado y aplicando propiedades de esperanza condicional:

$$Var(Y|X) = E[Y^2|X] - E[Y|X]^2$$

18.1. Propiedad (Pitagoras)

$$Var(Y) = E[Var(Y|X)] + Var(E[Y|X])$$