COMPILANDO CONOCIMIENTO

Probabilidad y Estadística

MATEMÁTICAS ESTADÍSTICAS

Oscar Andrés Rosas Hernandez

Febrero 2018

Índice general

Parte I

Probabilidad Clásica

Capítulo 1

Introducción

1.1. Notación

1.1.1. Experimento ε

Decimos que un experimento en probabilidad es cualquier proceso del cual se desconoce con determinación el resultado final.

Generalmente lo denotamos con mayúsculas.

1.1.2. Espacio Muestral S, Ω

Un espacio muestral asociado a un experimento es el conjunto de posibles resultados al momento de realizar el experimento.

Ejemplo:

Por ejemplo si ε_1 : Lanzar una moneda. Entonces tenemos que $S_1 = \{\ Cara, Cruz\ \}$

Si ε_2 : Lanzar un dado. Entonces tenemos que $S_2 = \{\ 1, 2, 3, 4, 5, 6\ \}$

1.1.3. Evento A

Un evento es simplemente algún subconjunto del espacio muestral.

Ejemplo:

Por ejemplo si ε_1 : Lanzar una moneda. Entonces tenemos que un evento puede ser $A_1 = \{\ Cara\ \}$

Si ε_2 : Lanzar un dado. Entonces tenemos que un evento puede ser $A_2 = \{\ 1, 2, 4\ \}, A_2.1 = \{\ 5\ \}$

1.2. Probabilidad $P(A)$

Definimos la probabilidad de un evento A como:

$$P(A) = \frac{|A|}{|S|}$$ Recuerda que A es un evento y S es espacio muestral

Creo que es muy obvio por la manera en que definimos a la probabilidad de un evento es un número real entre 0 y 1.

Por lo tanto:

- **NO hay probabilidades negativas**
- **NO hay probabilidades mayores a uno**

Entonces podemos reducir el problema de encontrar la probabilidad de un evento simplemente a dos partes:

- Encontrar la cardinalidad de dicho evento
- Encontrar la cardinalidad del espacio muestral de un experimento

1.2.1. Consecuencias

- $P(S) = 1$
- Si $A_1, A_2, \dots, A_n$ son eventos mutuamente excluyentes entonces:

$$P\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- Si $A_1, A_2, \dots$ son una colección infinita, pero contable de eventos mutuamente excluyentes entonces:

$$P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

- $P(\emptyset) = 0$
- $P(A') = 1 - P(A)$
- Si $A \subseteq B$ entonces $P(A) \leq P(B)$
- La probabilidad de la unión de n eventos de puede escribir de manera general como:

$$P\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i A_j) + \sum_{i<j<k}^{n} P(A_i A_j A_k) + \cdots + (-1)^{n+1} P\left(\cap_{i=1}^{n} A_i\right)$$

- Por consecuencia del caso general tenemos que:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Por consecuencia del caso general tenemos que:

$$\begin{aligned} P(A \cup B \cup C) = & \\ & P(A) + P(B) + P(C) \\ & - (P(A \cap B) + P(A \cap C) + P(B \cap C)) \\ & + P(A \cap B \cap C) \end{aligned}$$

1.3. Probabilidad Condicional

La probabilidad de que ocurra el Evento A conociendo que ya paso el Evento B se denota y define como:

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

Nota que para que todo esto tenga sentido $P(A) \neq 0$

Podemos notar entonces que el evento B tiene muchas interpretaciones como:

- La condición que ya esta dada
- Evento que se sabe que ya ocurrió o que es seguro que ocurra
- Espacio Muestral Reducido

1.3.1. Regla de la Multiplicación

Nota que $P(A \cap B)$ se puede ver como:

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

Podemos entonce redefinir la propabilidad condicional como:

$$\begin{aligned}
P(A|B) &= \frac{P(A \cap B)}{P(B)} && \text{Por definición} \\
&= \frac{P(B \mid A)P(A)}{P(B)} && \text{Por regla de multiplicación en el numerador} \\
&= \frac{P(B \mid A)P(A)}{P((B \cap A) \cup (B \cap A'))} && \text{Unión de eventos disconjuntos} \\
&= \frac{P(B \mid A)P(A)}{P(B \cap A) + P(B \cap A')} && \text{Aplicando regla de la multiplicación a cada uno} \\
&= \frac{P(B \mid A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}
\end{aligned}$$

1.4. Eventos Independientes

Dados 2 eventos que A, B son Independientes si y solo si $P(A) = P(A|B)$ y se escribe: $A \perp B$.

1.4.1. Propiedades

- Si $A \perp B$ entonces $P(A \cap B) = P(A)P(B)$

  Demostración:

  Si $A \perp B$ entonces $P(B) = P(B|A)$, por lo tanto $P(B) = \frac{P(B \cap A)}{P(A)}$

  Y solo despejas

- Si $A \perp B$ entonces $A' \perp B'$

  Demostración:

  Esta es clave:

$$\begin{aligned} P(A' \cap B') &= P((A \cup B)') \\ &= 1 - [P(A) + P(B) + P(A \cap B)] \\ &= 1 - [P(A) + P(B) + P(A \cap B)] \\ &= (1 - P(A))(1 - P(B)) \\ &= P(A')P(B') \end{aligned}$$

- Si $A \perp B$ entonces $P(A \cap B) \neq 0$ o la contrapositiva que dice que si son disjuntos entonces no son independientes

1.5. Teorema de Bayes

Considera un conjunto de eventos $\{ A_1, \ldots, A_n \}$ mutuamente exclusivos y tales que $\cup_{i=1}^{n} A_i = S$, es decir son particiones de S.

Entonces podemos escribir la propabilidad de un evento B como:

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

Gracias a esto podemos decir que:

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i \cap B)}{P(B)} \\ &= \frac{P(B|A_i)P(A_i)}{\displaystyle\sum_{i=1}^{n} P(A_i)P(B|A_i)} \end{aligned}$$

Capítulo 2

Variables Aleatorias Discretas

2.1. Variables Aleatorias

Una variable aleatoria es un función que asigna a cada elemento, en el espacio muestral a un número real.

Esta se denota con mayúsculas y no es un número, es una función. Para poner a un valor posible de una variable aleatoria lo denotamos con minúsculas.

2.1.1. Variables Aleatorias Discretas

Las variables aleatorias cuyo conjunto de valores posibles es finito o infinito contable entonces decimos que es una variable aleatoria discreta.

Ejemplo:

Por ejemplo considera que vas a lanzar 3 monedas, entonces tenemos que:

$$S = \{ \; ccc, ccx, cxc, xcc, xxc, xcx, cxx, xxx \; \}$$

Entonces podemos tener una variable aleatoria como:

Sea $X =$ Número de caras en 3 lanzamientos.

Entonces podemos decir que:

$$X(ccc) = 3, X(ccx) = 2, X(cxc) = 2, X(xcc) = 2, X(xxc) = 1, X(xcx) = 1, X(cxx) = 1, X(xxx) = 0$$

Por lo tanto los vales posibles son $0, 1, 2, 3$.

2.2. Función Probabilidad f_X

Es una función que toma todos los posibles valores una variable aleatoria y nos regresa un número real entre el 0 y el 1. Es decir:

$$f_X(x) = P(X = x)$$

Y tiene que cumplir que la suma de todos los posibles valores de la variable aleatoria den uno.

Más formalmente tenemos que la función probabilidad es aquella función que cumple que:

- $\forall a \in \mathbb{R}\ f_X(a) \geq 0$
- $\{\ x \ \mid\ f_X(x) \neq 0\ \}$ es un conjunto finito o numerable
- $\sum_x f_X(x) = 1$

Ejemplo:

Por ejemplo podemos definir la probabilidad del ejemplo pasado podemos definir la función $f_X(x) = \binom{3}{3} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{3-x}$ para $x \in \{\ 0, 1, 2, 3\ \}$.

Entonces tenemos que:

- La probabilidad de que $X = 0$ (caigan 0 caras) es $\frac{1}{8}$
- La probabilidad de que $X = 1$ (caigan 1 caras) es $\frac{3}{8}$
- La probabilidad de que $X = 2$ (caigan 2 caras) es $\frac{3}{8}$
- La probabilidad de que $X = 3$ (caigan 3 caras) es $\frac{1}{8}$

2.3. Función Probabilidad Acumulada F_X

La describimos como:

$$\overline{F}_X(x) = \sum_{i<x} f_X(i)$$

Podemos ver a la acumulada como una función fundamental, tal que podemos escribir a todas las demás:

- $P(X < x) = \overline{F}_X(x-1)$
- $P(X \leq x) = \overline{F}_X(x)$
- $P(X > x) = 1 - \overline{F}_X(x)$
- $P(X \geq x) = 1 - \overline{F}_X(x-1)$
- $P(a \leq X \leq b) = \overline{F}_X(b) - \overline{F}_X(a-1)$
- $P(a < X \leq b) = \overline{F}_X(b) - \overline{F}_X(a)$
- $P(a \leq X < b) = \overline{F}_X(b-1) - \overline{F}_X(a)$
- $P(a < X < b) = \overline{F}_X(b-1) - \overline{F}_X(a-1)$

2.4. Esperanza o Media

Decimos que el valor esperado o esperanza media de la variable X se define como:

$$\mu_X = E(X) = \sum_x x f_X(x) = \sum_x x P(X = x)$$

Representa un promedio ponderado de los valores posibles de la variable basado en sus probabilidades.

Es decir, si se repitiera el experimiento muchísimas veces el promedio de los resultados se iría aproximando a la media.

2.4.1. Propiedades

- Si X puede tomar un número infinito de valores entonces la esperanza de X existe si y solo si $\sum_x |x| f_X(x) < \infty$
- $E(g(x)) = \sum_x g(x) f_X(x)$
- Es un Operador Lineal, es decir:

  $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$
- Si X, Y son independientes entonces $E(XY) = E(X) + E(Y)$
- $E(a) = a$

2.5. Varianza

Decimos que la varianza de la variable X con $f_X(x)$ se define como:

$$v(X) = E((X - \mu)^2)$$

Pero usando esta definición podemos dar muchas otras definiciones utiles:

$$\begin{aligned} v(X) &= E((X - \mu)^2) \\ &= \sum_x (x - \mu)^2 P(X = x) \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

Podemos decir por su misma definición que la varianza siempre es positiva.

Es decir, este valor nos indica que tan lejos estan en promedio los valores de su misma media, es decir, que tan dispersa o concentrada esta la distribución de los datos.

2.5.1. Propiedades

- $V(a) = 0$

  Demostración:

  Sea $a = g(X)$, entonces su $\mu = a$, por lo tanto

  $$V(a) = E(a - a)^2) = E(0)^2) = 0 \tag{2.1}$$

- $v(aX) = a^2 v(X)$

  Demostración:

  $$\begin{aligned} V(aX) &= E(aX^2) - E^2(aX) &(2.2) \\ &= a^2 E(X^2) - a^2 E^2(X) &(2.3) \\ &= a^2 [E(X^2) - E^2(X)] &(2.4) \\ &= a^2 [v(X] &(2.5) \end{aligned}$$

- $v(X + Y) = v(X) + v(Y) + 2Cov(X, Y)$
- $v(X - Y) = v(X) - v(Y) - 2Cov(X, Y)$
- Si X y Y son independientes, entonces $v(X + Y) = v(X) + v(Y)$
- En general si $X_1, X_2, \ldots, X_n$ son variables aleatorias, entonces tenemos que:

  $$v\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} X_i + 2\sum_{i<j}^{n} Cov(X_i, X_j)$$

2.6. Desvianción Estandar

Decimos que la varianza de la variable X con $f_X(x)$ se define como:

$$\sigma(X) = \sqrt{v(X)}$$

Podemos decir por su misma definición que la varianza siempre es positiva.

Es decir, este valor nos indica que tan lejos estan en promedio los valores de su misma media, es decir, que tan dispersa o concentrada esta la distribución de los datos.

2.7. Covarianza

La definimos como:

$$Cov(X, Y) = E\left((X - \mu_X)(Y - \mu_Y)\right)$$

Es una manera de medir la dispersión conjunta de ambas variables.

2.7.1. Propiedades

- $Cov(X, Y) = E(XY) - \mu_X \mu_Y$

2.8. Momentos Centrales

Si X es una variable aleatoria tal que $E(X) = \mu_x$ entonces tenemos que:

- El k-ésimo momento central esta definida como:

$$\mu_k^c = E\left[(X-\mu)^k\right] = \sum_x (x-\mu)^k P(X=x) \tag{2.6}$$

- El k-ésimo momento alrededor del origen esta definida como:

$$\mu_k^0 = E\left[(X)^k\right] = \sum_x x^k P(X=x) \tag{2.7}$$

2.8.1. Propiedades

- Ya hemos trabajo con momentos, por ejemplo, la varianza la podemos definir como: $v(X) = E(X^2) - \mu^2$ donde $E(X^2)$ es un segundo momento central

  O bien, verla como $v(X) = E[(X-\mu)^2]$ es decir la varianza es el segundo momento central.

2.8.2. Función Generadora de Momentos

La podemos definir como:

$$\Psi_X(t) = E(e^{tX}) = \sum_x e^{tx} P(X = x)$$

Solemos decir que la k-ésima derivada de la función generadora de momentos evaluada en $t = 0$ da como resultado el k-ésimo momento central al origen

Es decir, siguen el siguiente patrón:

- $\Psi'_X(t = 0) = E(X)$
- $\Psi''_X(t = 0) = E(X^2)$
- $\Psi'''_X(t = 0) = E(X^3)$
- $\Psi^{(n)}_X(t = 0) = E(X^n)$

2.8.3. Propiedades

- Si $Y = aX + b$ entonces tenemos que:
  $\Psi_Y(t) = e^{bt}\Psi_X(at)$
- Nota que $\Psi_x(a) = E(e^{aX})$
- $\Psi_X(t = 0) = \mu_k$
- $\Psi_X(t = 0) = E(X)$
- Nota que si tuvieramos un montón de variables aleatorias $X_1, X_2, X_3, \dots, X_n$ Y decimos que $Y = \sum_{i=1}^{n} X_i$

  Entonces tenemos que $\Psi_Y(t) = \prod_{i=1}^{n} \Psi_{X_i}(t)$

2.9. Distribuciones Famosas

2.9.1. Bernoulli

Suponte un experimento en el que solo tienes dos salidas, $0, 1$, suponte que la probabilidad de que salga 0 es q y la que salga 1 es p.

Ahora veamos algunas caracteristicas:

- **Función Probabilidad**

  Esta esta fácil $f_X(x) = p^x q^{1-x}$

- **Esperanza de X**

  Nota que por definición tenemos que:

$$\begin{aligned}
\mu_X &= \sum_x xP(X = x) \\
&= 0(q) + 1(p) \\
&= p
\end{aligned}$$

- **Varianza**

  Por definición tenemos que:

$$\begin{aligned}
v(X) &= E(x^2) - \mu^2 \\
&= E(x^2) - p^2 \\
&= \sum_x x^2 P(X = x) - p^2 \\
&= 0(q) + 1(p) - p^2 \\
&= p - p^2 \\
&= p(1 - p) \\
&= pq
\end{aligned}$$

- **Acumulada**

  Por definición tenemos que:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ q & 0 \leq x < 1 \\ p & 1 \leq x \end{cases}$$

- **Función Genera de Momentos**

  Por definición tenemos que:

$$\begin{aligned}
\Psi(t) &= E(e^t X) \\
&= \sum_x e^t x P(X = x) \\
&= e^{0t}(q) + e^{1t}(p) \\
&= (q) + e^t(p)
\end{aligned}$$

2.9.2. Binomial

Es un experimento binomial que consiste en n ensayos Bernoulli independientees, con una probabilidad de exito individual constante.

La variable discreta es el número de exitos de los experimentos individuales donde sus posibles valores fueron $X = \{ 0, 1, \dots, n \}$.

Observa que la variable aleatoria se puede ver como la suma de Bernoullis independientes.

Ahora veamos algunas caracteristicas:

- **Función Probabilidad**

  Esta esta fácil $f_X(x) = \binom{n}{x} p^x q^{n-x}$

  ¿Porque? Porque estamos hablando de eventos independientes por lo tanto su probabilidad conjunta es el producto de las individuales, y literlalmente estamos usando la definición de combinación.

- **Función Acumulada**

  Esta esta fácil $F_X(x) = \sum_{i=0}^{x} \binom{n}{i} p^i q^{n-i}$

- **Esperanza de X**

  Podemos usar que la variable aleatoria se puede ver como la suma de Bernoullis independientes.

  Es decir $Y = \sum_x x_i$

$$
\begin{aligned}
\mu_X &= E(Y) \\
&= E(\sum_x x_i) \\
&= \sum_x E(x_i) \\
&= np
\end{aligned}
$$

- **Varianza**

  Por definición tenemos que:

  $$v(X) = npq$$

- **Función Genera de Momentos**

  Por definición tenemos que:

  $$\Psi(t) = \left(q + e^t p\right)^n$$

2.9.3. Geométrica

Supón que se repiten de manera independiente ensayos Bernoulli con una probabilidad de exito constante de p hasta obtener el primer exito.

La variable discreta es el número de experimentos hasta un primer exito donde sus posibles valores fueron $X = \{ 0, 1, \dots, \}$.

Observa que la variable aleatoria se puede ver como la suma de Bernoullis independientes.

Ahora veamos algunas caracteristicas:

- **Función Probabilidad**

  Esta esta fácil $f_X(x) = q^{x-1}p$

  Es decir, es la propabilidad de que todos los anteriores sean fracasos y el actual sea el exito.

- **Función Acumulada**

  $F_X(x) = 1 - q^x$

  Esta esta fácil:

$$\begin{aligned} F_X(x) &= \sum_{i=1}^{x} pq^{i-1} \\ &= p \sum_{i=0}^{x-1} q^i \\ &= p\frac{1-q^x}{1-q} \\ &= p\frac{1-q^x}{p} \\ &= 1 - q^x \end{aligned}$$

- **Esperanza de X**

  Esta es muy famosa: $E(X) = \frac{1}{p}$

  Demostración:

$$
\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} xP(X = x) \\
&= \sum_{x=1}^{\infty} xq^{x-1}p \\
&= p\sum_{x=1}^{\infty} xq^{x-1} \\
&= p\frac{d}{dx}\sum_{x=1}^{\infty} xq^{x-1} \\
&= p\frac{d}{dx}\sum_{x=1}^{\infty} q^{x} \\
&= p\frac{d}{dx}\frac{1}{1-q} \\
&= p\frac{1}{(1-q)^2} \\
&= p\frac{1}{p^2} \\
&= \frac{1}{p}
\end{aligned}
$$

- **Varianza**

  $v(X) = \frac{1-p}{p^2}$

$$\begin{aligned}
v(X) &= E(X^2) - \mu^2 \\
&= E(X^2) - \frac{1}{p^2} \\
&= \sum_{x=1}^{\infty} x^2 P(X = x) - \frac{1}{p^2} \\
&= \sum_{x=1}^{\infty} x^2 p q^{x-1} - \frac{1}{p^2} \\
&= p \sum_{x=1}^{\infty} x^2 q^{x-1} - \frac{1}{p^2} \\
&= p \frac{1+q}{(1-q)^3} - \frac{1}{p^2} \\
&= \frac{1+q}{p^2} - \frac{1}{p^2} \\
&= \frac{q}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}$$

- **Función Genera de Momentos**

  Por definición tenemos que:

$$\begin{aligned}
\Psi(t) &= E\left(e^{tX}\right) \\
\Psi(t) &= \sum_{x=1}^{\infty} e^{tx} q^{x-1} p \\
\Psi(t) &= p \sum_{x=1}^{\infty} e^{tx} q^{x-1} \\
\Psi(t) &= p \sum_{x=1}^{\infty} e^{tx} q^{x-1} \\
\Psi(t) &= pe^t \sum_{x=1}^{\infty} (qe^t)^x \\
\Psi(t) &= pe^t \frac{1}{1-qe^t} \\
\Psi(t) &= \frac{pe^t}{1-qe^t}
\end{aligned}$$

Capítulo 3

Combinatoria

3.1. Ideas Clave

3.1.1. Ordén vs Sin Ordén

En las muestras que estan ordenadas entonces el ordén de los elementos importa por ejemplo en los dígitos de un teléfono o en las letras de una palabra.

En las muestras que no estan ordenadas el ordén es irrelevante, por ejemplo en los elementos de un conjunto.

3.1.2. Remplazar vs No Remplazar

Las muestras con remplazo entonces estan permitidas, por ejemplo los números de la licencia.

Cuando la repetición no esta permitida, por ejemplo en un conjunto de números de lotería

3.2. Fórmulas

- Número de **permutaciones** de un conjunto de n objetos

$$n! \tag{3.1}$$

- Número de **muestras ordenadas** de tamaño r **con remplazo** de un conjunto de n objetos

$$n^r \tag{3.2}$$

- Número de **muestras ordenadas** de tamaño r **sin remplazo** de un conjunto de n objetos

$${}_nP_r = \frac{n!}{(n-r)!} = (n)(n-1)(n-2)\dots(n-r+1) \tag{3.3}$$

- Número de **muestras no ordenadas** de tamaño r **sin remplazo** de un conjunto de n objetos

  Esto es lo mismo que el número de subconjuntos de cardinalidad r de un conjunto de n elementos

$$\binom{n}{r} = {}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!} = \frac{(n)(n-1)(n-2)\dots(n-r+1)}{r!} \tag{3.4}$$

- Número de subconjuntos de un conjunto de n elementos:

$$2^n \tag{3.5}$$

3.3. Coheficientes Binomiales

3.3.1. Interpretación

Podemos interpretarlos de muchas maneras, algunas son:

- El número de formas de seleccionar k objetos de un conjunto de n elementos.
- El número de subconjuntos de elementos k en un conjunto de n elementos.

3.3.2. Propieades

- Número de **permutaciones** de un conjunto de n objetos

$$\binom{n}{r} = \frac{n!}{k!(n-k)!} \qquad \text{donde } n \in \mathbb{N} \text{ y } k \in \{ 1, \dots, n \} \tag{3.6}$$

$$= \frac{(n)(n-1)(n-2)\dots(n-k+1)}{k!} \tag{3.7}$$

- Número de **permutaciones** de un conjunto de n objetos

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{3.8}$$

- Propiedades Simetrícas

$$\binom{n}{k} = \binom{n}{n-k} \tag{3.9}$$

- Casos Especiales

$$\binom{n}{0} = \binom{n}{n} = 1 \qquad \binom{n}{1} = \binom{n}{n-1} = n \tag{3.10}$$

- Teorema del Binomio

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \tag{3.11}$$

- Teorema del Binomio (Caso Especial)

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1 \tag{3.12}$$

3.4. Propiedades

- Las formas de permutar n elementos en un círculo es $(n-1)!$

Bibliografía

[1] Leticia Cañedo Suárez *Probabilidad.* ESCOM, 2018