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:

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

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. Propiedades

- $P(S) = 1$
- Si $A_1, A_2, \ldots, A_n$ son eventos mutuamente excluyentes entonces:
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$
- $P(\emptyset) = 0$
- $P(A') = 1 - P(A)$
- 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:

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

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

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 $P(B) \neq 0$

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. Propiedades

- **Conservamos Propiedades**

  La propiedad condicional cumple las propiedades que ya vimos de una propiedad de un evento cualquiera, pero ahora el espacio muestral que antes era S se ha reducido.

  - $P\left(A \mid B\right) + P\left(A' \mid B\right) = 1$
  - $P\left(A \cup B \mid C\right) = P\left(A \mid C\right) + P\left(B \mid C\right) - P\left(A \cap B \mid C\right)$

- **Definición Alterna**

  Podemos redefinir a la probabilidad condicional como: $P\left(A \mid B\right) = \dfrac{|A \cap B|}{|B|}$

  Demostración:

  Esta es sencilla, muy sencilla:

  $$\begin{aligned} P\left(A \mid B\right) &= \frac{P(A \cap B)}{P(B)} && \text{Por definición de Condicional} \\ &= \frac{\frac{|A \cap B|}{|S|}}{\frac{|B|}{|S|}} && \text{Por definición de Probabilidad} \\ &= \frac{|A \cap B|}{|B|} && \text{Magia} \end{aligned}$$

- **Regla de Multiplicación**

  Podemos escribir a $P(A \cap B)$ en terminos de probabilidad condicional.

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

  Demostración:

  Mira: S $P(A|B) = \frac{P(A \cap B)}{P(B)}$ entonces $P(A|B)P(B) = P(A \cap B)$

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

1.4. Eventos Independientes

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

Es decir la ocurrencia de B no influye en nada a la ocurriencia de A, osea que pase o no pase B, a A le da igual.

1.4.1. Propiedades

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

  Demostración:

  Si $A \perp B$ entonces $B \perp A$ 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 $P(A \cap B) \neq 0$
- Si $P(A \cap B) = 0$ entonces A, B no son eventos independientes
- $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

1.4.2. Teorema de Bayes

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

Entonces podemos escribir la propabilidad de un evento B donde $B \subset S$ como:

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

Gracias a esto podemos decir que:

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

Capítulo 2

Combinatoria

2.1. Ideas Clave

2.1.1. 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.

2.1.2. 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

2.2. Permutación

Una permutación es un arreglo de objetos donde el ordén es importante.

Entonces definimos a ${}_nP_r$ a la cantidad de muestras ordenadas de tamaño r sin remplazo de un conjunto de n objetos.

Entonces decimos que:

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

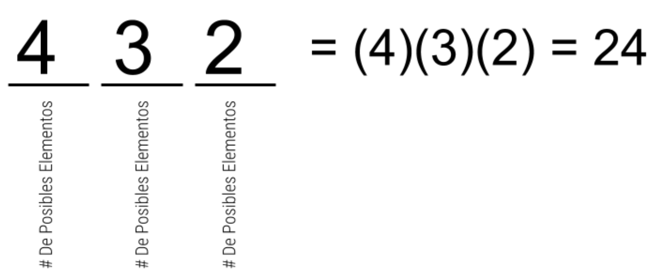
2.2.1. Ejemplos

Ejemplo 1:

Considera $S = \{\ a, b, c, d\ \}$, entonces podemos decir que:

- Hay 4 permutaciones distintas tomando solo una letra a la vez
- Hay 12 permutaciones distintas tomando solo dos letra a la vez
- Hay 24 permutaciones distintas tomando solo tres letra a la vez

Estas se pueden sacar facilmente con esta idea que creo que a todos nos enseñan, por ejemplo veamos como hacer el último punto:



2.3. Combinación

Una permutación es un arreglo de objetos donde el ordén NO es importante.

Entonces definimos a ${}_nC_r$ a la cantidad de muestras sin ordenadas de tamaño r sin remplazo de un conjunto de n objetos.

Entonces decimos que:

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

Esto tiene mucho sentido si lo ves desde otro angulo, pues en cuanto a las permutaciones tendremos $(n)(n-1)(n-2)\dots(n-r+1)$, pero resulta que muchas de esas permutaciones son basicamente la misma, solo cambiando el orden, así que si el orden ya no importa, es tan sencillo como dividir entre la cantidad de veces que podemos ordenar esas permutaciones de tamaño r

2.3.1. Combinaciones y Subconjuntos

Resulta ser que hay dos grande problemas clásicos de teoría de conjuntos que podemos resolver con combinaciones:

- El número de subconjuntos de cardinalidad r de un conjunto de n elementos

$$\binom{n}{r}$$

- Número de subconjuntos de un conjunto de n elementos:

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n$$

2.3.2. Ejemplos

Ejemplo 1:

Cuantos equipos se puede formar que incluyan 2 físicos y 1 matemático si se sabe que hay 4 físicos y 3 matemáticos.

Ya que no nos importa el orden esto esta mas sencillo de lo que parece:

$$\binom{4}{2}\binom{3}{1} = \frac{4!}{2!(4-2)!}\frac{3!}{1!(3-1)!} = 18$$

2.3.3. Propiedades Coheficientes Binomiales

- Propiedades Simetrícas

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

- Casos Especiales

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

- Teorema del Binomio

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

- Teorema del Binomio (Caso Especial)

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

Parte II

Variables Aleatorias Discretas

Capítulo 3

Variables Aleatorias Discretas

3.1. Variables Aleatorias

Una variable aleatoria es una función que asigna a cada elemento $E_i \in S$ en el espacio muestral a un número real $X(E_i) \in \mathbb{R}$, es decir, en español, lo que hace es que es una función que nos da información de una característica de cada elemento del espacio muestral.

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.

3.1.1. 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:

$$\begin{aligned} X(ccc) &= 3 \\ X(ccx) &= 2 \\ X(cxc) &= 2 \\ X(xcc) &= 2 \\ X(xxc) &= 1 \\ X(xcx) &= 1 \\ X(cxx) &= 1 \\ X(xxx) &= 0 \end{aligned}$$

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

3.2. Función Probabilidad f_X

3.2.1. Definición

También se le conoce como función de probabilidad puntual. Es una función que toma todos los posibles valores una variable aleatoria y nos regresa un número real entre el 0 y el 1 dado por la probabilidad de el valor de la variable aleatoria sea x. Es decir:

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

3.2.2. Propiedades

Es una función de probabilidad, es decir 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:

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

3.2.3. Ejemplos

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 $\frac{1}{8}$

3.3. Función P. Acumulada F_X

3.3.1. Definición

Describimos a la función de probabilidad acumulada como:

$$F_X(x) = \sum_{i \leq x} f_X(i)$$

3.3.2. Propiedades

- Una característica muy común es que:
  - $\lim_{x \to -\infty} F_X(x) = 0$
  - $\lim_{x \to \infty} F_X(x) = 1$
- Si $x_1 \leq x_2$ entonces $F_X(x_1) \leq F_X(x_2)$

3.3.3. Función Fundametal

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

- $P(X = x) = F_X(x) - F_X(x-1)$
- $P(X < x) = F_X(x-1)$
- $P(X \leq x) = F_X(x)$
- $P(X > x) = 1 - F_X(x)$
- $P(X \geq x) = 1 - F_X(x-1)$
- $P(a \leq X \leq b) = F_X(b) - F_X(a-1)$
- $P(a < X \leq b) = F_X(b) - F_X(a)$
- $P(a \leq X < b) = F_X(b-1) - F_X(a)$
- $P(a < X < b) = F_X(b-1) - F_X(a-1)$

3.4. Esperanza o Media

3.4.1. Definición

Decimos que el valor esperado, esperanza ó 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.

3.4.2. Propiedades

- Si X puede tomar un número infinito de valores entonces la esperanza de X existe si y solo si $\sum_x |x| f_X(x) < \infty$
- Podemos dar una definición al evaular la esperanza sobre una función:

  $$E(g(x)) = \sum_x g(x)\ f_X(x)$$

- Es un Operador Lineal, es decir:

  $$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

- Si X, Y son independientes entonces:

  $$E(XY) = E(X)E(Y)$$

- Si a es una constante, entonces:

  $$E(a) = a$$

3.5. Varianza

3.5.1. Definición

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

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

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.

3.5.2. Desviación Estandar

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

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

Se usa generalmente por las unidades que tiene la varianza, nada mas

3.5.3. Propiedades

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

  Demostración:

  Esto esta demasiada sencillo:

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

- $V(a) = 0$

  Demostración:

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

$$\begin{aligned} V(a) &= E(a-a)^2) \\ &= E(0)^2) \\ &= 0 \end{aligned}$$

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

  Demostración:

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

- $v(X+Y) = v(X) + v(Y) + 2Cov(X,Y)$

  Demostración:

  Esta esta larga:

$$\begin{aligned} v(X+Y) &= E(X+Y)^2 - (E(X+Y))^2 \\ &= E(X) + E(Y) + 2E(XY) - (E(X+Y))^2 \\ &= E(X) + E(Y) - (E(X) + E(Y))^2 + 2E(XY) \\ &= E(X) + E(Y) - E^2(X) - E^2(Y) + 2E(XY) - 2E(X)E(Y) \\ &= E(X) - E^2(X) + E(Y) - E^2(Y) + 2E(XY) - 2E(X)E(Y) \\ &= v(x) + v(Y) + 2E(XY) - 2E(X)E(Y) \\ &= v(x) + v(Y) + 2Cov(X,Y) \end{aligned}$$

- $v(X - Y) = v(X) - v(Y) - 2Cov(X, Y)$

  Demostración:

  Es lo mismo, que la de arriba :v

- Si X y Y son independientes, entonces $v(X + Y) = v(X) + v(Y)$

  Demostración:

  Es lo mismo que la de arriba, solo recuerda que si X, Y son independientes entonces tenemos que $Cov(X, Y) = 0$

- En general si $X_1, X_2, \ldots, X_n$ son variables aleatorias, entonces tenemos que:

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

  Demostración:

  Es inducción :v

3.6. Covarianza

3.6.1. Definición

Sea X, Y dos variables independientes, entonces definimos a la covarianza como:

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

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

3.6.2. Propiedades

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

  Demostración:

  Podemos demostrar esto, bien facil:

  $$\begin{aligned} Cov(X,Y) &= E\left((X-\mu_X)(Y-\mu_Y)\right) \\ &= E\left(XY - X\mu_Y - \mu_X Y - \mu_X\mu_Y\right) \\ &= E(XY) - E(X\mu_Y) - E(\mu_X Y) + E(\mu_X\mu_Y) \\ &= E(XY) - \mu_Y E(X) - \mu_X E(Y) + E(\mu_X\mu_Y) \\ &= E(XY) - \mu_Y\mu_X - \mu_X\mu_Y + E(\mu_X\mu_Y) \\ &= E(XY) - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y \\ &= E(XY) - \mu_Y\mu_X \end{aligned}$$

- La covarianza de 2 variables independientes es cero

  Demostración:

  Mira esto:

  $$\begin{aligned} Cov(X,Y) &= E(XY) - \mu_Y\mu_X && \text{Si es que son independientes} \\ &= E(X)E(Y) - \mu_Y\mu_X && \text{Recuerda que } E(X) = \mu_X \\ &= 0 \end{aligned}$$

3.7. 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:

$$\mu_k^c = E\left[(X-\mu)^k\right] = \sum_x (x-\mu)^k P(X=x) \tag{3.1}$$

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

$$\mu_k^0 = E\left(X^k\right) = \sum_x x^k P(X=x) \tag{3.2}$$

3.7.1. Propiedades

- Ya hemos trabajo con momentos, veamos algunos:
    - **La Esperanza**
      Esta se puede ver como el primer momento alrededor del origen
      $\mu_X = \mu_1^0$
    - **La Varianza**
      Esta se puede ver como el primer momento alrededor central

      $$v(X) = E[(X-\mu)^2] = \mu_2^c$$

      O bien podemos verlo como $v(X) = E(X^2) - \mu^2$ donde $E(X^2)$ es un segundo momento alrededor del origen

3.7.2. 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:

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

3.7.3. Propiedades

- Nota que $\Psi_x(a) = E(e^{aX})$
- Si $Y = aX + b$ entonces tenemos que:
  $\Psi_Y(t) = e^{bt}\Psi_X(at)$

  Demostración:

  Esta esta fácil, sea $Y = aX + b$:

$$\begin{aligned}\Psi_Y(t) &= E(e^{tY}) \\ &= E(e^{t(aX+b)}) \\ &= E(e^{atX+tb}) \\ &= E(e^{atX}e^{tb}) \\ &= e^{tb}E(e^{atX}) \\ &= e^{tb}\Psi_X(at)\end{aligned}$$

- $\Psi_X(t = 0) = \mu_k$
- $\Psi_X(t = 0) = E(X)$
- Nota que si tuvieramos un montón de variables aleatorias $X_1, X_2, X_3, \dots, X_n$ Y decimos que $Y = \displaystyle\sum_{i=1}^{n} X_i$

  Entonces tenemos que:

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

  Demostración:

  Esta también es importante:

$$\begin{aligned}\Psi_Y(t) &= E\left(e^{tY}\right) \\ &= E\left(e^{t\sum_{i=1}^{n} X_i}\right) \\ &= E\left(\prod_{i=1}^{n} e^{X_i}\right) \\ &= \prod_{i=1}^{n} E\left(e^{X_i}\right) \\ &= \prod_{i=1}^{n} \Psi_{X_i}(t)\end{aligned}$$

Capítulo 4

Distribuciones Famosas

4.1. Bernoulli

4.1.1. Definición

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

Variable Aleatoria

Ahora, nuestra variable aleatoria sera:

X : Resultado del experimento

Ahora, los valores posibles que puede tomar son muy muy sencillos, basicamente porque solo hay 2 opciones, o salio bien, o salio mal, es decir X puede tomar los valores $0, 1$.

Solemos entonces escribir esta distribución como:

$X \sim Ber(x; p)$

donde:

- x: Nuestra variable
- p: Probabilidad de Exito

4.1.2. Función Probabilidad

Esta es clásica:

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

Demostración:

Esta esta fácil, podemos hacerla por partes, como parece más natural y ver que:

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

Podemos extender esta idea de muchas maneras a una expresión, la que damos es solo una de ellas.

4.1.3. Función P. Acumulada

Por definición tenemos que:

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

Y esta conviene mejor dejarla así.

4.1.4. Esperanza

Esta es la distribución con la esperanza más fácil que veras:

$$E(X) = p$$

Demostración:

Nota que por definición tenemos que:

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

4.1.5. Varianza

Esta es igualmente sencilla:

$$v(X) = p(1 - p) = p - p^2 = pq$$

Demostración:

Por definición tenemos que:

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

4.1.6. Función Generadora

Esta igual es muy bonita:

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

Demostración:

Por definición tenemos que:

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

4.2. Binomial

4.2.1. Definición

Un experimento binomial consiste en n ensayos Bernoulli independientes, con una probabilidad de exito individual constante e igual en todos los experimentos

La variable aleatoria discreta nos medirá 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, es decir $X = X_1 + X_2 + X_3 + \cdots + X_n$

Variable Aleatoria

Ahora, nuestra variable aleatoria sera:

X : Número de Exitos en los experimentos

Ahora, los valores posibles que puede tomar son muy muy sencillos, suponte que haremos n experimentos, entonces X tiene que tomar valores entre $0 \dots n$

Solemos entonces escribir esta distribución como:

$X \sim Bin(x; n, p)$

donde:

- x: Nuestra variable
- p: Probabilidad de Exito de cada experimiento Bernoulli

Definición Alterna

Sea X_i variables aleatorias de Bernoulli, tal que esten definidas por:

$$X_i = \begin{cases} 1 & \text{Si el i ésimo experimento fue exitoso} \\ 0 & \text{Si el i ésimo experimento fue un fallo} \end{cases} \qquad \forall i \in [1, 2, \dots, n]$$

Nota que cada X_i es independiente

Entonces podemos ver a una Binomial como suma de Bernoulli INDEPENDIENTES es decir:

$X = X_1 + X_2 + \cdots + X_n$

4.2.2. Función Probabilidad

Esta esta fácil:

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

Demostración:

¿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.

4.2.3. Función P. Acumulada

Esta esta fácil:

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

Demostración:

Literalmente es la definición.

4.2.4. Esperanza

Esta también es sencilla:

$$E(X) = np$$

Demostración:

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

Es decir $X = \sum_x X_i$

$$\begin{aligned}
\mu_X &= E(X) \\
&= E(\sum_x X_i) \\
&= \sum_x E(X_i) \\
&= np
\end{aligned}$$

4.2.5. Varianza

Esta es bonita también:

$$v(X) = npq$$

Demostración:

Esta la haremos usando propiedades de la varianza, recuerda que son la suma de eventos independientes:

$$\begin{aligned} v(X) &= v\left(\sum_{i=0}^{n} X_i\right) \\ &= \sum_{i=1}^{n} v(X_i) \\ &= \sum_{i=1}^{n} pq \\ &= npq \end{aligned}$$

4.2.6. Función Generadora

Por definición tenemos que:

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

Demostración:

Espera, espera, me explico mejor, lo que pasa es la binomial se puede ver como una suma de variables independientes entonces solo basta con recordar que ya demostramos que la función generadora de momentos de una suma de variables independientes es el producto de la función generadora de momentos de cada una.

Es decir:

$$\begin{aligned} \Psi_X(t) &= \Psi_{\sum_{i=1}^{n} X_i}(t) \\ &= \prod_{i=1}^{n} \Psi_{X_i}(t) \\ &= \prod_{i=1}^{n} q + e^t p \\ &= \left(q + e^t p\right)^n \end{aligned}$$

4.2.7. Propiedades

- Sean $X_1, X_2, \dots, X_k$ k variables aleatorias discretas independientes y $X_i \sim Bin(x_i; n_i, p)$ con $i = 1, 2, \dots, k$.

  Entonces la variable aleatoria $X = X_1 + X_2 + \cdots + X_k$ tiene una distribución binomial tal que $X \sim Bin\left(x; \sum_{i=1}^{k} n_i, p\right)$

  Demostración:

  Veamos que:

$$\begin{aligned}
\Psi_X(t) &= \Psi_{X_1+X_2+\cdots+X_k}(t) \\
&= \prod_{i=1}^{k} \Psi_{X_i} \\
&= (q+pe^t)^{n_1}(q+pe^t)^{n_2}\dots(q+pe^t)^{n_k} \\
&= (q+pe^t)^{\sum_{i=i}^{k} n_i}
\end{aligned}$$

  Es decir $\Psi_X(t) = (q+pe^t)^n$ con $n = \sum_{i=i}^{k} n_i$, es decir vimos que $X \sim Bin\left(x; \sum_{i=1}^{k} n_i, p\right)$

4.3. Geométrica

4.3.1. Definición

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 son $X = \{ 1, 2, 3, \dots \}$.

Variable Aleatoria

Ahora, nuestra variable aleatoria sera:

X : Número de experimentos hasta un exito

Ahora, los valores posibles que puede tomar son muy muy sencillos, porque puede que pase en el primer experimento o en el segundo, es decir la variable aleatoria puede tomar cualquier valor en los enteros positivos.

Solemos entonces escribir esta distribución como:

$$X \sim G(x; p)$$

donde:

- x: Nuestra variable
- p: Probabilidad de Exito de cada experimento Bernoulli

4.3.2. 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.

4.3.3. Función P. Acumulada

Esta también es sencilla:

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

Demostración:

Esta esta fácil, solo espero que recuerdes la suma de la serie geométrica:

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

4.3.4. Esperanza

Esta es muy famosa, y lo repito, que no se te olvide la geométrica

$$E(X) = \frac{1}{p}$$

Demostración:

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

4.3.5. Varianza

Esta también es sencilla:

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

Demostración:

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

4.3.6. Función Generadora

Esta también saldra por definición no somos cobardes:

$$\Psi(t) = \frac{pe^t}{1 - qe^t}$$

Demostración:

Por definición tenemos que:

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

4.3.7. Propiedades

- Sea $X \sim G(x; p)$ entonces tenemos que $P(X > a) = q^a$ con a un natural positivo

  Demostración:

$$
\begin{aligned}
P(Y > a) &= 1 - F_X(a) && \text{Definición de Acumulada} \\
&= 1 - (1 - q^a) && \text{Talacha} \\
&= q^a && \text{Bingo}
\end{aligned}
$$

- Podemos decir que esta distribución no tiene memoria es decir que ya voy a experimentos y no ha pasado nada entonces, debería ya ser mas probable que para $a + b$ ya me tocará, me explico mejor ... ¿No? ¡Pues no!

  Espera, me explico mejor:

  Sea $X \sim G(x; p)$ entonces $P(X > a + b \mid X > a) = P(X > b)$ con a, b un naturales positivos

  Demostración:

$$
\begin{aligned}
P(X > a + b \mid X > a) &= \frac{P(X > a + b \text{ y } X > a)}{P(X > a)} && \text{Definición de Condicional} \\
&= \frac{P(X > a + b)}{P(X > a)} && \text{Sentido común} \\
&= \frac{p^{a+b}}{p^a} && \text{Teorema pasado} \\
&= p^{(a+b)-a} && \text{Exponentes} \\
&= p^b && \text{Usando teorema pasado} \\
&= P(X > b)
\end{aligned}
$$

4.4. Hiper-Geométrica

4.4.1. Definición

Supongamos que tenemos una población de tamaño r, ahora esta particionado de 2 maneras, con elementos del tipo r_1 o r_2, consideraremos exitosos los elementos de tipo r_1

Ahora vamos a tomar una muestra aleatoria de tamaño n sin reemplazo ni sustitución.

Ahora, solo por notación si es que n, es decir la muestra es menor que nuestra población r la llamamos muestra, pero si $n = r$ entonces decimos que es un censo.

Variable Aleatoria

Ahora, nuestra variable aleatoria sera:

X : Número de elementos del tipo r_1 en una muestra aleatoria de tamaño n

Ahora, los valores posibles que puede tomar son $max(0, n - r_2) \leq x \leq min(r_1, n)$

Ahora, ¿Porque esos números tan feos?

Por un lado tenemos que que lo peor que nos puede pasar son dos cosas:

- O bien si el tamaño de tu muestra n es muy pequeña, entonces puedes tener toda la mala suerte del mundo y que pase que todas caigan donde tu no querías por lo tanto, podría pasar que no seleccionará ningún elemento de r_1 entonces $X = 0$.
- Pero, pero que pasaría que tu n fuera lo suficientemente grande tal que incluso si seleccionará todos los que no quería r_2, aún quedarán elementos por seleccionar, entonces los mínimos elementos r_1 que podría seleccionar es $n - r_2$, entonces $X = n - r_2$

El cálculo para el límite mayor es parecido

- Si tienes que $n <= r_1$, entonces lo mejor que te puede pasar es que todos caigan donde tu quieras
- Si es que $n > r_1$ entonces puedes seleccionar todos los elementos de r_1

Solemos entonces escribir esta distribución como:

$$X \sim H(x; r_1, r_2, n)$$

donde:

- x: Nuestra variable
- r_1: Tamaño de nuestra población de interes
- r_2: Tamaño de la población que no es de interes, sale de $r_2 = r - r_1$
- n: Tamaño de la muestra

Definición Alterna

Sea X_i variables aleatorias de Bernoulli, tal que esten definidas por:

$$X_i = \begin{cases} 1 & \text{Al sacar el elemento i-esímo fue de tipo } r_1 \\ 0 & \text{Al sacar el elemento i-esímo NO fue de tipo } r_1 \end{cases} \qquad \forall i \in [1, 2, \dots, n]$$

Entonces podemos ver a una hipergeométrica como:

$$X = X_1 + X_2 + \cdots + X_n$$

Lo importante aquí es que como cada una de ellas no son independientes

4.4.2. Función Probabilidad

Esta esta muy interesante, es:

$$f_X(x) = \frac{\binom{r_1}{x}\binom{r_2}{n-x}}{\binom{r}{n}}$$

Idea de la Demostración:

Antes que nada, mira lo bonito que sale, arriba tienes r_1, r_2 y $r_1 + r_2 = r$ y abajo tienes que $n - x, x$ y $n - x + x = n$

Ahora, abajo estamos colocando las posibles formas de escojer conjunto de n elementos de una espacio de r elementos, y arriba es la probabilidad conjunta de que primero tengamos x elementos de r_1 y $n - x$ elementos de r_2

4.4.3. Función P. Acumulada

Esta esta muy interesante, es:

$$F_X(x) = P(X \leq x) = \frac{1}{\binom{r}{n}} \sum_{i=max(0,n-r_2)}^{x} \binom{r_1}{i}\binom{r_2}{n-i}$$

Idea de la Demostración:

Es por definición men :v

4.4.4. Esperanza

Esta esta muy interesante:

$$E(X) = n\left(\frac{r_1}{r}\right)$$

Demostración:

Primero que nada porque usando la definición o algo así nos van a salir cosas horribles, así que mejor empecemos por otro lado.

Sea X_i variables aleatorias de Bernoulli, tal que esten definidas por:

$$X_i = \begin{cases} 1 & \text{Al sacar el elemento i-esímo fue de tipo } r_1 \\ 0 & \text{Al sacar el elemento i-esímo NO fue de tipo } r_1 \end{cases}$$

Ahora, como son variables de Bernoulli, podemos encontrar bien facil su esperanza como $E(X_i) = p = \frac{r_1}{r}$

Ahora, como ya te esperabas, nota que nuestra variable aleatoria hipergeometrica es la suma de las otras $X = \sum_{i=1}^{n} X_i$

Ahora como la esperanza es un bonito operador lineal tenemos que:

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} E(X_i) \\ &= \sum_{i=1}^{n} \frac{r_1}{r} \\ &= n\left(\frac{r_1}{r}\right) \end{aligned}$$

4.4.5. Varianza

Esta esta muy dificíl, así que se deja para el lector :p

$$v(X) = n\left(\frac{r_1}{r}\right)\left(1 - \frac{r_1}{r}\right)\left(1 - \frac{n-1}{r-1}\right)$$

4.4.6. Función Generadora

Esta esta muy interesante, es:

$$\Psi_X(t) = E(e^{tX}) = \frac{1}{\binom{r}{n}} \sum_{i=max(0,n-r_2)}^{min(n,r_1)} e^{ti} \binom{r_1}{i} \binom{r_2}{n-i}$$

4.4.7. Relación con la Binomial

Suponte que tienes una población de r elementos, con r_1 de un tipo "bueno" y $r_2 := n-r$ de un tipo "malo".

Ahora vamos a tomar una muestra aleatoria de tamaño n, sea entonces X : Número de elementos de tipo bueno en nuestra muetra

ahora podemos tener entonces 2 posibles distribuciones dependiendo de una pregunta clave.

¿Hay reemplazo?

- **Si es que tiene Reemplazo:**

  Entonces lo que pasa es que la probabilidad de exito es constante, por lo tanto es simplemente n experimentos de tipo Bernoulli.

  La probabilidad de exito también es bastante sencilla de sacarse, es como $p = \frac{r_1}{r}$

  Por lo tanto podemos decir que:

  $$X \sim B(x; n, p)$$

  Es decir, tendremos (como ya sabemos como se comporta una binomial):

  - $E(X) = np = \frac{nr_1}{r}$
  - $v(X) = npq = \frac{nr_1}{r}\left(1 - \frac{r_1}{r}\right)$

- **Si es que NO tiene Reemplazo:**

  Entonces lo que pasa es que la probabilidad de exito ya no es constante, de hecho, llegamos a la definición de la Hipergeométrica, es decir:

  Por lo tanto podemos decir que:

  $$X \sim H(x; n, r_1, r_2)$$

  Es decir, tendremos (como ya sabemos como se comporta una hipergeometrica):

  - $E(X) = np = \frac{nr_1}{r}$
  - $v(X) = npq = \frac{nr_1}{r}\left(1 - \frac{r_1}{r}\right)\left(1 - \frac{n-1}{r-1}\right)$

Es decir, sin importar si hay o no reemplazo, el valor esperado es el mismo, pero si es que no hay reemplazos tenemos una varianza es menor.

Podemos también darnos cuenta de que mientras más población la diferencia entre el reemplazo y sin reemplazo cada vez será menor.

4.5. Poisson

Variable Aleatoria

La variable aleatoria que cuenta el número de ocurrencías en un periodo de tiempo o espacio físico dado se le llama Poisson.

X : Número de ocurrencias en un periodo de tiempo o espacio dado

Ahora, los valores posibles que puede tomar es $0, 1, 2, \ldots$

Solemos entonces escribir esta distribución como:

$X \sim P(x; \lambda)$

donde:

- x: Nuestra variable
- λ: Es el número promedio de ocurrencias en el periodo o espacio dado

4.5.1. Función Probabilidad

Esta esta muy interesante, es:

$$f_X(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

4.5.2. Función P. Acumulada

Esta esta muy interesante, es:

$$F_X(x) = P(X \leq x) = e^{-\lambda} \sum_{i=0}^{x} \frac{\lambda^i}{i!}$$

Idea de la Demostración:

Es por definición men :v

4.5.3. Función Generadora

Esta esta muy interesante, es:

$$\Psi_X(t) = e^{\lambda(e^t - 1)}$$

Demostración:

Usando la función probabilidad puntual tenemos que:

$$\begin{aligned}
\Psi_X(t) &= E(e^{tX}) \\
&= \sum_{x=0}^{\infty} e^{tx} P(X = x) \\
&= \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t\lambda)^x}{x!} \\
&= e^{-\lambda} e^{e^t\lambda} \\
&= e^{\lambda(e^t - 1)}
\end{aligned}$$

4.5.4. Esperanza

Esta esta muy interesante:

$$E(X) = \lambda$$

Demostración:

Esta sale o bien por definición o usando la generadora de momentos y diciendo que:

$$\begin{aligned}\Psi(t=0)' &= e^{\lambda(e^t-1)}\lambda(e^t)\Big|_{t=0}\\ &= e^0\lambda(e^0)\\ &= \lambda\end{aligned}$$

4.5.5. Varianza

Esta esta muy interesante:

$$v(X) = \lambda$$

Demostración:

Esta sale o bien por definición o usando la generadora de momentos y dicieno que:

$$\begin{aligned}\Psi(t=0)'' &= e^{\lambda(e^t-1)}\lambda(e^t) + \lambda(e^t)^2 e^{\lambda(e^t-1)}\Big|_{t=0}\\ &= \lambda + \lambda^2\end{aligned}$$

Por lo tanto la varianza es $v(X) = \lambda + \lambda^2 - \lambda^2 = \lambda$

4.5.6. Relación con la Binomial

Considera $X \sim Bin(x; n, p)$, entonces cuando más grande sea n más nuestra binomial se va a parecer a una Poisson con $\lambda = np$

Demostración:

Considera $X \sim Bin(x; n, p)$, entonces esta más que claro que por ser una variable aleatoria que se distribuye con una binomial que $P(X = x) = f_X = \binom{n}{x} p^x q^{n-x}$.

Ahora como ya demostramos en las propiedades de la Poisson, vamos a suponer por un minuto que $\lambda = np$, es decir $p = \frac{\lambda}{n}$ entonces tenemos que:

$$P(X = x) = f_X = \binom{n}{x} p^x q^{n-x} = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Ahora veamos que es lo que pasa cuando tenemos una x muy muy grande, es decir:

$$\begin{aligned}
\lim_{n\to\infty} P(X = x) &= \lim_{n\to\infty} \binom{n}{x} p^x q^{n-x} \\
&= \lim_{n\to\infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \frac{\lambda^x}{x!} \frac{n!}{n^x(n-x)!} \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{n-x+1}{n} \dots \frac{n-1}{n} \frac{n}{n} \left(1 - \frac{\lambda}{n}\right)^{n} \frac{1}{(n-\lambda)^x} \\
&= \frac{\lambda^x}{x!} e^{-\lambda} \lim_{n\to\infty} \frac{n-x+1}{n} \dots \frac{n-1}{n} \frac{n}{n} \frac{1}{(n-\lambda)^x} \\
&= \frac{\lambda^x}{x!} e^{-\lambda} \lim_{n\to\infty} (1) \dots (1) \\
&= \frac{\lambda^x}{x!} e^{-\lambda}
\end{aligned}$$

Por lo tanto cuando más grande sea n más nuestra binomial se va a parecer a una Poisson con $\lambda = np$

4.5.7. Propiedades

- Sea $X_1, X_2, \dots, X_k$ variables aleatorias independientes y $X_i \sim P(x_i; \lambda_i)$
  Entonces $X = \sum_{i=1}^{k} X_i$ cumple que X se distribuye como una Poisson con parámetro $\lambda = \sum_{i=1}^{k} \lambda_i$

  Demostración:

  Sabes que para cada una de ellas tienes que $\Psi_{X_i}(t) = e^{\lambda_i(e^t-1)}$ entonces al ser una suma de variables aleatorias tenemos que su generadora es el producto de cada generadora, es decir:

$$
\begin{aligned}
\Psi_X(t) &= \prod i = 1^k \Psi_{X_i}(t) \\
&= \prod i = 1^k e^{\lambda_i(e^t-1)} \\
&= e^{(\sum_{i=1}^{k} \lambda_i)(e^t-1)}
\end{aligned}
$$

  Es decir se parece muchísimo a una Poisson con una lamda igual a $\lambda = \sum_{i=1}^{k} \lambda_i$

Parte III

Variables Aleatorias Continuas

Capítulo 5

Variables Aleatorias Continuas

5.1. Variables Aleatorias

Una variable aleatoria es una función que asigna a cada elemento $E_i \in S$ en el espacio muestral a un número real $X(E_i) \in \mathbb{R}$, es decir, en español, lo que hace es que es una función que nos da información de una característica de cada elemento del espacio muestral.

5.1.1. Variables Aleatorias Continuas

Las variables aleatorias cuyo conjunto de valores posibles el de los números reales, nos permiten medir un parámetro continuo.

Recuerda que los números reales son densos, eso quiere decir que entre cuales quiera dos reales, podemos encontrar otro real entre ambos.

Obviamente se conservan practicamente todas las propiedades de cuando trabajamos con las variables aleatorias discretas.

5.2. Función Probabilidad f_X

5.2.1. Definición

Vamos a definir a la función de probabilidad de una variable aleatoria continua como aquella función $f_X(x)$ para la cual siempre se cumplan 2 cosas:

- $\displaystyle\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
- $\displaystyle P(a < X < b) = \int_a^b f_X(x)\, dx$

Recuerda que las probabilidades en el caso continuo se puede ver como áreas bajo la curva delimitada según el interes.

5.2.2. Probabilidad Puntual

También se le conoce como función de probabilidad puntual. Es tecnicamente la misma que en las variables aleatorias discretas, pero al estar hablando de puede tomar cualquier real, entonces decimos que:

$$P(X = x) = 0$$

Esto nos lleva una propiedad muy importante que vamos a ocupar a cada rato:

$$P(a < X < b) = P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b)$$

Demostración:

- $P(a < X \leq b) = P(a < X < b) + P(b) = P(a < X < b) + 0 = P(a < X < b)$
- $P(a \leq X < b) = P(a < X < b) + P(a) = P(a < X < b) + 0 = P(a < X < b)$
- $P(a \leq X \leq b) = P(a < X < b) + P(a) + P(b) = P(a < X < b) + 0 + 0 = P(a < X < b)$

5.3. Función P. Acumulada F_X

5.3.1. Definición

Vamos a definir a la función de distribución o acumulada como:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x) dx$$

Recuerda que las probabilidades en el caso continuo se puede ver como áreas bajo la curva delimitada según el interes.

5.3.2. Propiedades

- $f_X(x) = \dfrac{d\ F_X(x)}{dx}$ si es que tiene sentido la derivada en ese punto sino $f_X(x) = 0$

5.4. Esperanza

5.4.1. Definición

Vamos a definir a la esperanza de una variable aleatoria continua como:

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

5.5. Varianza

5.5.1. Definición

Vamos a definir a la varianza de una variable aleatoria continua como:

$$\begin{aligned} v(x) &= \int_{-\infty}^{\infty} (x - \mu_x)^2\ f(x)\ dx \\ &= \left(\int_{-\infty}^{\infty} x^2 f(x)\ dx \right) - \mu_x^2 \end{aligned}$$

5.6. Función Generadora de Momentos

5.6.1. Definición

Vamos a definir a la esperanza de una variable aleatoria continua como:

$$\Psi_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\ dx$$

Capítulo 6

Distribuciones Continuas Famosas

6.1. Uniforme

6.1.1. Definición

La mas sencilla de todas las distribuciones continuas es la uniforme.

Variable Aleatoria

Ahora, los valores posibles que puede tomar nuestra variable aleatoria es $a < x < b$.
Solemos entonces escribir esta distribución como:

$$X \sim U(x; a, b)$$

donde:

- a: Inicio de la muestra
- b: Inicio de la muestra

6.1.2. Función Probabilidad

Esto es lo que distingue a una distribución uniforme:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{en otro caso} \end{cases}$$

Idea de la Demostración:

La razón de que sea así es que tiene que cumplir que $\int_{-\infty}^{\infty} f(x)dx = 1$ y claro que lo cumple:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_a^b \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \int_a^b dx \\ &= \frac{1}{b-a}[b-a] \\ &= 1 \end{aligned}$$

6.1.3. Función P. Acumulada

Esta esta muy interesante, es:

$$F_X(x) = \frac{x-a}{b-a}$$

Idea de la Demostración:

Es por definición men :v

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f(x)dx \\ &= \int_a^x \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \int_a^x dx \\ &= \frac{x-a}{b-a} \end{aligned}$$

6.1.4. Esperanza

Esta esta muy interesante:

$$E(X) = \frac{a+b}{2}$$

Demostración:

Esta también sale bonita:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f_X(x) dx \\
&= \int_a^b x \frac{1}{b-a} dx \\
&= \left(\frac{1}{b-a}\right) \int_a^b x dx \\
&= \left(\frac{1}{b-a}\right) \frac{b^2-a^2}{2} \\
&= \left(\frac{1}{b-a}\right) \frac{(b-a)(b+a)}{2} \\
&= \frac{a+b}{2}
\end{aligned}$$

6.1.5. Varianza

Esta esta muy dificíl, pero si se puede:

$$v(X) = \frac{(a-b)^2}{12}$$

Demostración:

Primero hay que hacer que:

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x)dx \\
&= \int_a^b x^2 \frac{1}{b-a} dx \\
&= \left(\frac{1}{b-a}\right) \int_a^b x^2 dx \\
&= \left(\frac{1}{b-a}\right) \frac{b^3-a^3}{3} \\
&= \left(\frac{1}{b-a}\right) \frac{(b-a)(b^2+ab+a^2)}{3} \\
&= \frac{b^2+ab+a^2}{3}
\end{aligned}$$

Ahora podemos sacar a la varianza:

$$\begin{aligned}
v(X) &= E(X^2) - \mu_X^2 \\
&= E(X^2) - \frac{(a+b)^2}{4} \\
&= \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4} \\
&= \frac{4b^2+4ab+4a^2}{12} - \frac{3a^2+6ab+3b^2}{12} \\
&= \frac{a^2+b^2-2ab}{12} \\
&= \frac{(a-b)^2}{12}
\end{aligned}$$

6.1.6. Función Generadora

Esta esta muy interesante, es:

$$\Psi_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$$

Demostración:

Ahora vamos a demostrarlo:

$$\begin{aligned}
E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\
&= \left(\frac{1}{b-a} \right) \int_a^b e^{tx} dx \\
&= \left(\frac{1}{b-a} \right) \frac{e^{bt} - e^{at}}{t} \\
&= \frac{e^{bt} - e^{at}}{t(b-a)}
\end{aligned}$$

6.2. Exponencial

6.2.1. Definición

Se utiliza generalmente para representar una distribución del tiempo que transcurre antes de la ocurrencia de un suceso.

Variable Aleatoria

Ahora, los valores posibles que puede tomar nuestra variable aleatoria es $x > 0$

Solemos entonces escribir esta distribución como:

$$X \sim E(x; \beta)$$

donde:

- β: Es un párametro que tiene que ser un entero positivo

6.2.2. Función Probabilidad

Esto es lo que distingue a una distribución uniforme:

$$f_X(x) = \begin{cases} \beta e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

6.2.3. Función P. Acumulada

Esta esta muy interesante, es:

$$F_X(x) = 1 - e^{-\beta x}$$

Idea de la Demostración:

Es por definición men :v

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f(t)dt \\ &= \frac{0}{x}\beta e^{-\beta t}dt \\ &= -e^{-\beta t}\Big|_{0}^{x} \\ &= 1 - e^{-\beta x} \end{aligned}$$

6.2.4. Función Generadora

Esta esta muy interesante, es:

$$\Psi_X(t) = \frac{\beta}{\beta - t}$$

Demostración:

Ahora vamos a demostrarlo:

$$\begin{aligned} E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\ &= \int_{0}^{\infty} e^{tx} \beta e^{-\beta x} dx \\ &= \beta \int_{0}^{\infty} e^{(t-\beta)x} dx \\ &= \beta \frac{1}{t-\beta} e^{(t-\beta)x} \Big|_0^{\infty} \\ &= \beta \frac{1}{t-\beta} (-1) \\ &= \frac{\beta}{\beta - t} \end{aligned}$$

6.2.5. Esperanza

Esta esta muy interesante:

$$E(X) = \frac{1}{\beta}$$

Demostración:

Esta también sale bonita:

$$\begin{aligned} E(X) &= \Psi'_X(0) \\ &= \frac{d}{dx} t \frac{\beta}{\beta - t} \Big|_0 \\ &= \frac{\beta}{(\beta - t)^2} \Big|_0 \\ &= \frac{\beta}{(\beta)^2} \\ &= \frac{1}{\beta} \end{aligned}$$

6.2.6. Varianza

Esta esta muy dificíl, pero si se puede:

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

Demostración:

$$\begin{aligned}
E(X^2) &= \Psi''_X(0) \\
&= \frac{d}{dx} \frac{\beta}{(\beta - t)^2} \Big|_0 \\
&= \frac{2\beta}{(\beta - t)^3} \Big|_0 \\
&= \frac{2\beta}{(\beta)^3} \\
&= \frac{2}{\beta^2}
\end{aligned}$$

Entonces la varianza sale rapido:

$$\begin{aligned}
v(X) &= E(X^2) - \mu_X^2 \\
&= \frac{2}{\beta^2} - \frac{1}{\beta^2} \\
&= \frac{1}{\beta^2}
\end{aligned}$$

6.2.7. Propiedades

- Podemos decir que esta distribución no tiene memoria, me explico mejor.

  Sea $P(X > a + b \mid X > a) = P(X > b)$ con a, b un naturales positivos

  Demostración:

$$\begin{aligned} P(X > a + b \mid X > a) &= \frac{P(X > a + b \text{ y } X > a)}{P(X > a)} \\ &= \frac{P(X > a + b)}{P(X > b)} \\ &= \frac{1 - F_X(a + b)}{1 - F_X(b)} \\ &= \frac{e^{-a\beta - b\beta}}{e^{-b\beta}} \\ &= e^{-b\beta} \end{aligned}$$

- Sea $X_1, X_2, \dots, X_k$ variables aleatorias independientes y $X_i \sim E(x_i; \beta)$

  Entonces $X = min(X_1, X_2, \dots, X_k)$ también cumple que X se distribuye como una Exponencial con parámetro $\beta = n\beta$

  Demostración:

  Esa sale rápido:

$$\begin{aligned} P(X > t) &= P(X_1 > t)P(X_2 > t) \dots (X_k > t) \\ &= e^{-t\beta} \dots e^{-t\beta} \\ &= e^{-t(n\beta)} \end{aligned}$$

  Es decir se parece muchísimo a una Exponencial con una beta igual a $\beta = n\beta$

- Sea $X_1, X_2, \dots, X_\alpha$ variables aleatorias independientes y $X_i \sim E(x_i; \beta)$

  Entonces $X = X_1 + X_2 + \cdots + X_\alpha$ se distruye como como $X \sim \Gamma(x; \alpha, \beta)$

  Demostración:

  Nota que:

$$\Psi_{X_i}(t) = \frac{\beta}{\beta - t}$$

  Entonces tenemos que:

$$\begin{aligned} \Psi_X(t) &= \prod_{i=1}^{\alpha} \frac{\beta}{\beta - t} \\ &= \left(\frac{\beta}{\beta - t} \right)^{\alpha} \end{aligned}$$

  Es decir, es una función generadora de la gamma, y ya :v

6.3. Gamma

6.3.1. Función Gamma

La función Gamma se puede ver para $a > 0$ como:

$$\Gamma(a) = \int_0^\infty x^{a-1}\ e^{-x} dx$$

Propiedades

La función gamma tiene un par de propieades interesantes

- $\Gamma(1) = 1$

  Demostración:

  Nota que:

$$\begin{aligned}
\Gamma(1) &= \int_0^\infty x^{1-1}\ e^{-x} dx \\
&= \int_0^\infty e^{-x} dx \\
&= -e^{-x} \Big|_0^\infty \\
&= e^0 - 0 \\
&= 1
\end{aligned}$$

- $\Gamma(a) = (a-1)\Gamma(a-1)$

  Demostración:

  Esto es clave:

$$\Gamma(a) = \int_0^\infty x^{a-1}\ e^{-x} dx$$

  Esta se puede intentar resolver por partes, entonces tenemos que:

  - $u = x^{a-1}$
  - $du = (a-1)x^{a-2} dx$
  - $dv = e^{-x}$
  - $v = -e^{-x}$

  Entonces decimos que:

$$\begin{aligned}
\Gamma(a) &= \int_0^\infty x^{a-1}\ e^{-x} dx \\
&= -x^{a-1} e^{-x} \Big|_0^\infty + \int_0^\infty e^{-x} (a-1) x^{a-2} dx \\
&= \Gamma(a) = (a-1)\Gamma(a-1)
\end{aligned}$$

- Si $n \in \mathbb{Z}^+$ entonces $\Gamma(n) = (n-1)!$

  Demostración:

  Es simplemente recursividad es decir:

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-2)\\ &= (n-1)(n-2)\Gamma(n-3)\\ &= \dots\\ &= (n-1)!\end{aligned}$$

6.3.2. Definición

Se utiliza generalmente para representar una distribución del tiempo que transcurre antes de la ocurrencia de un suceso.

Variable Aleatoria

Ahora, los valores posibles que puede tomar nuestra variable aleatoria es $x > 0$

Solemos entonces escribir esta distribución como:

$$X \sim \Gamma(x; a, \beta)$$

donde:

- a: Entero positivo
- β: Es un párametro que tiene que ser un entero positivo

6.3.3. Función Probabilidad

Esto es lo que distingue a una distribución uniforme:

$$f_X(x) = \begin{cases} \dfrac{\beta^a}{\Gamma(a)}\ x^{a-1}e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

6.3.4. Función Acumulada

Esta esta rara, pero ve que.

$$F_X(x) = 1 - F_X^P(\alpha - 1) \qquad \text{Con } \lambda = \beta x$$

Es decir:

$$F_X(x) = 1 - e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}$$

Demostración:

Esta esta bien genial: Primero veamos que lo que queremos sacar es:

$$P(X \geq x) = \frac{\beta^\alpha}{\Gamma(x)} \int_x^\infty x^{\alpha-1}e^{-\beta x}dx$$

Ahora hay que notar al patrón que sale de hacer por partes:

$$\begin{aligned} P(X \geq x) &= \frac{\beta^\alpha}{\Gamma(x)} \int_x^\infty x^{\alpha-1}e^{-\beta x}dx \\ &= \frac{(\beta x)^{\alpha-1}}{\Gamma(\alpha-1)!}e^{-\beta x} + \frac{(\beta x)^{\alpha-2}}{\Gamma(\alpha-2)!}e^{-\beta x} + \frac{(\beta x)^{\alpha-3}}{\Gamma(\alpha-3)!}e^{-\beta x} + \dots \end{aligned}$$

Ahora ve que α es un entero positivo

Por lo tanto tenemos que:

$$P(X \geq x) = \sum_{i=0}^{\alpha} -1 \frac{(\beta x)^i}{\Gamma(i)!}e^{-\beta x}$$

Nota que esto no es más que una Poisson, más explicitamente tenemos que:

$$\sum_{i=0}^{\alpha} -1 \frac{(\beta x)^i}{\Gamma(i)!}e^{-\beta x} = F_X^P(\alpha - 1)$$

Entonces nos damos cuenta al final que:

$$F_X(x) = 1 - P(X \geq x) = 1 - F_X^P(\alpha - 1)$$

Donde tenemos que $\alpha = \beta x$

6.3.5. Generadora de Momentos

Esto es lo que distingue a una distribución uniforme:

$$\Psi_X(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}$$

Demostración:

$$\begin{aligned}
\Psi_X(t) &= E(e^{tX}) \\
&= \frac{\beta^n}{\Gamma^n} \int_{-\infty}^{\infty} e^{tx} x^{\alpha-1} e^{-\beta x} dx \\
&= \frac{\beta^n}{\Gamma^n} \int_{-\infty}^{\infty} e^{-x(\beta-t)} x^{\alpha-1} dx
\end{aligned}$$

Ahora vemos que: $x = \frac{u}{\beta - t}$ $u = [\beta - t]x \to du = [\beta - t]dx \; to \; dx = \frac{du}{\beta - t}$

Ahora vemos que:

$$\begin{aligned}
\Psi_X(t) &= \frac{\beta^n}{\Gamma^n} \int_{-\infty}^{\infty} \left[\frac{u}{b - t}\right]^{\alpha-1} e^{-u} \frac{du}{\beta - t} \\
&= \frac{\beta^{\alpha}}{(\beta - t)^{\alpha}\Gamma(\alpha)} \int_0^{\infty} u^{\alpha-1} e^{-u} \\
&= \frac{\beta^{\alpha}}{(\beta - t)^{\alpha}} \\
&= \left(\frac{\beta}{\beta - t}\right)^{\alpha}
\end{aligned}$$

6.3.6. Esperanza

Usando la función generadora tenemos que:

$$E(X) = \frac{\alpha}{\beta}$$

Demostración:

$$\begin{aligned}
E(X) &= \Psi'_X(0) \\
&= \frac{d}{dx}t\left[\frac{\beta}{\beta - t}\right]^{\alpha} \Big|_{t=0} \\
&= \beta^{\alpha}\frac{d}{dx}t\,(\beta - t)^{-\alpha} \Big|_{t=0} \\
&= \beta^{\alpha}\alpha(\beta - t)^{-\alpha-1} \Big|_{t=0} \\
&= \beta^{\alpha}\alpha\beta^{-\alpha-1} \\
&= \frac{\alpha}{\beta}
\end{aligned}$$

6.3.7. Varianza

Usando la función generadora tenemos que:

$$v(X) = \frac{\alpha}{\beta^2}$$

Demostración:

$$\begin{aligned}
E(X^2) &= \Psi'_X(0) \\
&= \beta^{\alpha}\alpha\frac{d}{dx}t[\beta - t]^{-\alpha-1} \Big|_{t=0} \\
&= \beta^{\alpha}\alpha(-\alpha - 1)\frac{d}{dx}t[\beta - t]^{-\alpha-2}(-1) \Big|_{t=0} \\
&= \beta^{\alpha}\alpha(\alpha + 1)\frac{d}{dx}t[\beta - t]^{-\alpha-2} \Big|_{t=0} \\
&= \beta^{\alpha}\alpha(\alpha + 1)\beta^{-2} \\
&= \frac{\alpha(\alpha + 1)}{\beta^2}
\end{aligned}$$

Por lo tanto finalmente tenemos que:

$$\begin{aligned}
v(X) &= E(X^2) - E(X)^2 \\
&= \frac{\alpha(\alpha + 1)}{\beta^2} - \left(\frac{\alpha}{\beta}\right) \\
&= \frac{\alpha}{\beta^2}
\end{aligned}$$

6.3.8. Propiedades

- Si tenemos una distribución Gamma pero donde $\alpha = 1$, entonces lo que en realidad tenemos es una distribución exponencial con la misma β
- Sea $X_1, X_2, \dots, X_k$ variables aleatorias independientes y $X_i \sim \Gamma(x_i; \alpha_i, \beta)$
  Entonces $X = X_1 + X_2 + \cdots + X_k$ se distruye como como $X \sim \Gamma(x; \sum_{i=1}^{k} \alpha_i, \beta)$

  Demostración:

  Nota que:

$$\Psi_{X_i}(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}$$

  Entonces tenemos que:

$$\begin{aligned}\Psi_X(t) &= \prod_{i=1}^{k} \left(\frac{\beta}{\beta - t}\right)^{\alpha} \\ &= \left(\frac{\beta}{\beta - t}\right)^{\sum_{i=1}^{k} \alpha_i}\end{aligned}$$

  Es decir, es una función generadora de la gamma, y ya :v

6.4. Normal

6.4.1. Definición

Es la más famosa de todas las continuas, es simetrica con respecto al valor central μ

Variable Aleatoria

Ahora, los valores posibles que puede tomar nuestra variable aleatoria son todos los reales

Solemos entonces escribir esta distribución como:

$$X \sim N(x; \mu, \sigma^2)$$

donde:

- μ: El valor central, es cualquier real
- σ^2: La varianza, es un real positivo

Nota que si pasa que $X \sim N(x; \mu = 0, \sigma^2 = 1)$ entonces se suele denotar como: $Z \sim N(0, 1)$

Se le conoce como normal estandar.

6.4.2. Estandarización

Ya que tenemos una normal estandar, es decir una que cumple que $\mu = 0$, $\sigma^2 = 1$

Ahora, hay una forma muy facil de "estandarizar" es decir de transformar cualquier punto sobre una normal a ese punto en la estandar esta esta dada por:

$$Z = \frac{X - \mu}{\sigma}$$

Demostración:

Lo que hacemos es primero sacar la distancia al punto central, una vez que lo tengamos lo que hacemos es dividir entre la desviación estandar porque es una nueva escala una que nunca cambia.

Y obviamente tiene su inversa:

$$X = Z\sigma + \mu$$

6.4.3. Función Probabilidad

Esto es lo que distingue a una distribución uniforme:

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

Demostración que es una función de Probabilidad:

Supon que $I = \int_{-\infty}^{\infty} f_X(x)dx$, por lo tanto basta con hacer demostrar que $I^2 = 1$.

Primero haremos el cambio de variable a $u = \dfrac{x-\mu}{\sigma}$ y $du = \dfrac{dx}{\sigma}$, es decir, vamos a mejor resolver la estandar:

$$I = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$

Entonces decimos ahora que:

$$\begin{aligned}
I^2 &= \frac{1}{2\pi}\left(\int_{-\infty}^{\infty} e^{\frac{1}{2}u^2} du\right)^2 && (6.1)\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(u^2+v^2)} dudv && (6.2)
\end{aligned}$$

Ahora simplemente cambiamos a coordenada polares:

$$\begin{aligned}
I^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{1}{2}(u^2+v^2)} dudv && (6.3)\\
&= \frac{1}{2\pi}\int_{0}^{2\pi}\int_{0}^{\infty} e^{-\frac{r^2}{2}} rdrd\theta && (6.4)\\
&= \frac{1}{2\pi}\int_{0}^{2\pi}\left[-e^{-\frac{r^2}{2}}\Big|_0^{\pi}\right] d\theta && (6.5)\\
&= \frac{1}{2\pi}\int_{0}^{2\pi} d\theta && (6.6)\\
&= \frac{1}{2\pi}\theta\Big|_0^{2\pi} && (6.7)\\
&= 1 && (6.8)
\end{aligned}$$

6.4.4. Generadora de Momentos

Esta es sencilla:

$$\Psi_X(t) = e^{t\mu + \frac{1}{2}(t^2\sigma^2)}$$

Demostración:

Hagamos la definición:

$$\begin{aligned}
\Psi_X(t) &= \int_{-\infty}^{\infty} e^{tx} f_x dx && \text{Por definición} \\
&= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx && \text{Definición igual :v} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu+\sigma u) - \frac{1}{2}u^2} du && \text{La hacemos estandar } u = \frac{x-\mu}{\sigma} \text{ y } du = \frac{dx}{\sigma} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\mu} e^{-\frac{1}{2}\left(u^2 - 2t\sigma u\right)^2} du && \text{Ponemos bonito} \\
&= \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(u^2 - 2t\sigma u + (t^2\sigma^2) - (t^2\sigma^2)\right)^2} du && \text{Ponemos bonito} \\
&= \frac{e^{t\mu + \frac{1}{2}(t^2\sigma^2)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(u^2 - 2t\sigma u + (t^2\sigma^2)\right)^2} du && \text{Completamos el cuadrado} \\
&= \frac{e^{t\mu + \frac{1}{2}(t^2\sigma^2)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\ (z)^2} dz && z = u - t\sigma \ dz = du \\
&= e^{t\mu + \frac{1}{2}(t^2\sigma^2)} \int_{-\infty}^{\infty} \sqrt{2\pi} e^{-\frac{1}{2}\ (z)^2} dz && \text{Acomodo para hacer una normal estandar} \\
&= e^{t\mu + \frac{1}{2}(t^2\sigma^2)} && \text{Recuerda que por ser función de probabilidad}
\end{aligned}$$

6.4.5. Media

Mira, que loco :v

$$E(X) = \mu$$

Demostración:

Mira, que bonito es todo:

$$\begin{aligned} E(X) &= \Psi_X(0)' \\ &= e^{t\mu + \frac{1}{2}(t^2\sigma^2)} \left[\mu + \frac{2t\sigma^2}{2}\right] \Big|_0 \\ &= e^0(\mu) \\ &= \mu \end{aligned}$$

6.4.6. Varianza

Mira, que loco :v

$$v(X) = \sigma^2$$

Demostración:

Mira, que bonito es todo:

$$\begin{aligned} v(X) &= E(X^2) - \mu^2 \\ &= \Psi_X(0)'' - \mu^2 \\ &= \mu^2 + \sigma^2 - \mu^2 \\ &= \sigma^2 \end{aligned}$$

6.4.7. Propiedades

- Sea $X_1, X_2, \dots, X_k$ variables aleatorias independientes y $X_i \sim N(x_i; \mu_i, \sigma_i^2)$
  Entonces $X = X_1 + X_2 + \dots + X_k$ se distruye como como $X \sim N(x_i; \sum_{i=1}^k \mu_i, \sum_{i=1}^k \sigma_i^2)$

  Demostración:

  Empecemos por ver que $X_i \sim N(x; \mu_i, \sigma_i^2)$ Es decir $\Psi_{X_i}(t) = e^{t\mu_i + \dfrac{1}{2}(t^2\sigma_i^2)}$ entonces vemos que:

$$
\begin{aligned}
\Psi_X(t) &= \Psi_{\sum_{i=0}^n X_i}(t) \\
&= \prod_{i=1}^n \Psi_{X_i}(t) \\
&= e^{\sum_{i=1}^n t\mu_i + \dfrac{1}{2}(t^2\sigma_i^2)} \\
&= e^{t\left(\sum_{i=1}^n \mu_i\right) + \dfrac{1}{2}\left(t^2\left(\sum_{i=1}^n \sigma_i^2\right)\right)}
\end{aligned}
$$

  Por lo tanto como puedes ver tiene una media de $\mu_X = \sum_{i=1}^n \mu_i$ y una varianza de $\sigma_X^2 = \sum_{i=1}^n \sigma_i^2$

- Sea $X_i \sim N(x_i; \mu, \sigma^2)$
  Entonces $Y = aX + b$ se distruye como como $Y \sim N(y; a\mu + b, a\sigma^2)$

  Demostración:

  Empecemos por ver que $X \sim N(x; \mu, \sigma^2)$ entonces vemos que:

$$
\begin{aligned}
\Psi_Y(t) &= E(e^{tY}) \\
&= E(e^{t(aX+b)}) \\
&= e^{tb}E(e^{taX}) \\
&= e^{tb}\Psi_X(at) \\
&= e^{tb}e^{at\mu + \frac{1}{2}a^2t^2\sigma^2} \\
&= e^{t(b+a\mu) + \frac{1}{2}a^2t^2\sigma^2}
\end{aligned}
$$

  Por lo tanto como puedes ver tiene una media de $\mu_Y = a\mu + b$ y una varianza de $\sigma_Y^2 = a^2\sigma^2$

- Sea $X_1, X_2, \dots, X_k$ variables aleatorias independientes y $X_i \sim N(x_i; \mu, \sigma^2)$
  Entonces $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ se distruye como como $\overline{X} \sim N(\overline{x}; \mu, \frac{\sigma^2}{n})$

  Demostración:

  Colorario jovén, es un colorario :v

Capítulo 7

Probabilidad Hardcore

7.1. Teorema Central del Límite

7.1.1. Definición

Garantiza que para n (tamaño de muestra) suficientemente grande ($n \geq 30$) entonces:

- Si tienes una variable aleatoria que es la media muestral de otras variables aleatorias entonces mientras mas variables aleatoria muestrees entonces más se va aproximar:

  $\overline{X} \approx N(x; \mu = \mu, \sigma^2 = \dfrac{\sigma^2}{n})$

- Si tienes una variable aleatoria que esta compuesta por la suma de muchas otras variables aleatorias entonces mientras mas grande sea la cantidad de variables aleatorias más se aproxima a una normal:

  $\sum_{i=0}^{n} X_i \approx N(x; \mu_x = n\mu, \sigma_x^2 = n\sigma^2)$

7.2. Teorema Central de Chebyshue

7.2.1. Definición

Si no se conoce la distribución de la variable aleatoria, pero por circunstancias raras tenegamos la media y la varianza entonces podemos conocer información de la variable.

Tenemos que:

$$
\begin{aligned}
P(\mu - k\sigma \leq X \leq \mu + k\sigma) &\geq 1 - \frac{1}{k^2} \\
P(|X - \mu| \leq k\sigma) &\geq 1 - \frac{1}{k^2} \\
P(|X - \mu| > k\sigma) &< \frac{1}{k^2}
\end{aligned}
$$

Parte IV

Probabilidades Conjuntas

Capítulo 8

Variables A. Multidimensionales

8.1. Función de Probabilidad

La función de probabilidad conjunta de una variable bidimensional se puede ver de dos maneras:

8.1.1. Discreto

La definimos como:

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

Y cumple con que:

- $P(X = x, Y = y) \in [0, 1]$
- $\displaystyle\sum_y \sum_x P(X = x, Y = y) = 1$

8.1.2. Continuo

Es la función de dos variables que cumple que:

- $f_{X,Y}(x,y)$ nunca es negativa
- $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1$

8.2. Distribución Acumulada

8.2.1. Discreto

La definimos como:

$$F_{X,Y} := P(X \leq x, Y \leq y)$$

8.2.2. Continuo

La definimos como:

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

Nota que $\dfrac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = f_{X,Y}(x,y)$

8.3. Distribución Marginal

8.3.1. Discreto

La definimos de 2 maneras:

- $f_X(x) := \sum_{y} P(X = x, Y = y)$
- $f_Y(y) := \sum_{x} P(X = x, Y = y)$

8.3.2. Continuo

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

8.4. Distribución Condicional

8.4.1. Discreto

Podemos definirlo como:

$$f_{X|y}(x|y) = \frac{P(X = x, Y = y)}{P(Y = y)} \quad \forall x$$

Sus valores posibles son los valores de la variable aleatoria X.

Tenemos una distribución diferente para cada valor de Y

8.4.2. Continuo

Podemos definirlo como:

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

Sus valores posibles son los valores de la variable aleatoria X.

Tenemos una distribución diferente para cada valor de Y

8.5. Independencia de Variables

Sea una variable aleatoria dimensional X, Y, son independientes si y solo si:

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

O dicho en otras palabras:

$$f_{X,Y}(x, y) = f_X(x)F_Y(y)$$

8.5.1. Propiedades

- X, Y son linealmente independientes si y solo si $P(X = x|Y = y) = P(X = x)$ o dicho de otra manera si $f_{X|Y}(x, y) = f_X(x)$

Parte V

CheatSheet - Formulario

Capítulo 9

CheatSheet - Formulario

9.1. Teoría de Conjuntos

Nombre	Propiedad
Operaciones Básicas	
Complemento	$A' = \{ \ x \ \mid \ x \notin X \ \}$
Intersección	$A \cap B = \{ \ x \ \mid \ x \in A \text{ y } x \in B \ \}$
Unión	$A \cup B = \{ \ x \ \mid \ x \in A \text{ ó } x \in B \ \}$
Leyes de Morgan	
Morgan sobre Unión	$(A \cup B)' = A' \cap B'$
Morgan sobre Intersección	$(A \cap B)' = A' \cup B'$
Combinatoria	
Complemento	$A' = \{ \ x \ \mid \ x \notin X \ \}$
Intersección	$A \cap B = \{ \ x \ \mid \ x \in A \text{ y } x \in B \ \}$
Unión	$A \cup B = \{ \ x \ \mid \ x \in A \text{ ó } x \in B \ \}$

9.2. Combinatoria

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

  $n!$

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

  n^r

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

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

- 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

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

- Número de subconjuntos de un conjunto de n elementos:

  2^n

  item Las formas de permutar n elementos en un círculo es:

  $(n-1)!$

9.2.1. Propiedades Coheficientes Binomiales

- Propiedades Simetrícas

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

- Casos Especiales

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

- Teorema del Binomio

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

- Teorema del Binomio (Caso Especial)

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

9.3. Probabilidad Básica

Definimos la probabilidad de un evento A como:

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

9.3.1. Propiedades

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

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

- $P(\emptyset) = 0$
- $P(A') = 1 - P(A)$
- 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:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_iA_j) + \sum_{i<j<k}^{n} P(A_iA_jA_k) + \cdots + (-1)^{n+1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

- 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:

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

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

9.4. Probabilidad Condicional

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

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

9.4.1. Propiedades

- **Conservamos Propiedades**

  La propiedad condicional cumple las propiedades que ya vimos de una propiedad de un evento cualquiera, pero ahora el espacio muestral que antes era S se ha reducido.

  - $P(A \mid B) + P(A' \mid B) = 1$
  - $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$

- **Definición Alterna**

  Podemos redefinir a la probabilidad condicional como: $P(A \mid B) = \dfrac{|A \cap B|}{|B|}$

- **Regla de Multiplicación**

  Podemos escribir a $P(A \cap B)$ en terminos de probabilidad condicional.

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

9.5. Eventos Independientes

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

9.5.1. Propiedades

- Si $A \perp B$ entonces $P(A \cap B) = P(A)P(B)$
- Si $A \perp B$ entonces $A' \perp B'$
- Si $A \perp B$ entonces $P(A \cap B) \neq 0$
- Si $P(A \cap B) = 0$ entonces A, B no son eventos independientes

9.5.2. Teorema de Bayes

Considera un conjunto de eventos $\{ A_1, \dots, A_n \}$ mutuamente excluyentes y tales que $\displaystyle\bigcup_{i=1}^{n} A_i = S$, es decir son particiones de S.
Entonces podemos escribir la propabilidad de un evento B donde $B \subset S$ como:

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

Gracias a esto podemos decir que:

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

9.6. Variables Aleatorias Discretas

9.6.1. Función Probabilidad f_X

Propieadades

Es una función de probabilidad, es decir 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:

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

9.6.2. Función P. Acumulada F_X

Definición

Describimos a la función de probabilidad acumulada como:

$$F_X(x) = \sum_{i \leq x} f_X(i)$$

Propiedades

- Una característica muy común es que:
  - $\lim_{x \to -\infty} F_X(x) = 0$
  - $\lim_{x \to \infty} F_X(x) = 1$
- Si $x_1 \leq x_2$ entonces $F_X(x_1) \leq F_X(x_2)$

Función Fundametal

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

- $P(X = x) = F_X(x) - F_X(x-1)$
- $P(X < x) = F_X(x-1)$
- $P(X \leq x) = F_X(x)$
- $P(X > x) = 1 - F_X(x)$
- $P(X \geq x) = 1 - F_X(x-1)$
- $P(a \leq X \leq b) = F_X(b) - F_X(a-1)$
- $P(a < X \leq b) = F_X(b) - F_X(a)$
- $P(a \leq X < b) = F_X(b-1) - F_X(a)$
- $P(a < X < b) = F_X(b-1) - F_X(a-1)$

9.6.3. Esperanza o Media

Definición

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

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

Propiedades

- Si X puede tomar un número infinito de valores entonces la esperanza de X existe si y solo si $\sum_x |x| f_X(x) < \infty$
- Podemos dar una definición al evaular la esperanza sobre una función:

  $$E(g(x)) = \sum_x g(x)\ f_X(x)$$

- Es un Operador Lineal, es decir:

  $$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

- Si X, Y son independientes entonces:

  $$E(XY) = E(X)E(Y)$$

- Si a es una constante, entonces:

  $$E(a) = a$$

9.6.4. Varianza

Definición

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

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

Desvianción Estandar

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

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

Se usa generalmente por las unidades que tiene la varianza, nada mas

Propiedades

- $v(X) = E(X^2) - (E(X))^2$
- $V(a) = 0$
- $v(aX) = a^2 v(X)$
- $v(X+Y) = v(X) + v(Y) + 2Cov(X, Y)$
- $v(X-Y) = v(X) - v(Y) - 2Cov(X, Y)$
- Si X y Y son independientes, entonces $v(X+Y) = v(X) + v(Y)$
- En general si $X_1, X_2, \dots, X_n$ son variables aleatorias, entonces tenemos que:

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

9.6.5. Covarianza

Definición

Sea X, Y dos variables independientes, entonces definimos a la covarianza como:

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

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

Propiedades

- $Cov(X, Y) = E(XY) - \mu_X \mu_Y$
- La covarianza de 2 variables independientes es cero

9.6.6. 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:

$$\mu_k^c = E\left[(X-\mu)^k\right] = \sum_x (x-\mu)^k P(X=x) \tag{9.1}$$

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

$$\mu_k^0 = E\left(X^k\right) = \sum_x x^k P(X=x) \tag{9.2}$$

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:

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

Propiedades

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

  Entonces tenemos que:

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

Bibliografía

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