MEFE 2025: Comentarios sobre distribuciones

Matías Leizerovich *

April 12, 2025

En este apunte haremos algunos comentarios sobre la guía 2 y 3, a partir de algunas discusiones que surgieron sobre distribuiones discretas y continuas.

1 Distribuciones discretas

Relación entre la Distribución Binomial y la Binomial negativa La distribución binomial negativa se utiliza para modelar el número de ensayos necesarios n hasta obtener un número fijo de éxitos k. En este caso, la distribución binomial se utiliza para modelar el número de éxitos en un número fijo de ensayos. En la distribución binomial negativa el número de ensayos no es fijo, sino que se detiene cuando se alcanza un número determinado de éxitos.

Relación entre la Distribución Binomial y la Hipergeométrica La distribución binomial se utiliza cuando se realizan ensayos independientes con reposición, mientras que la distribución hipergeométrica se utiliza cuando se realizan ensayos sin reposición. En el caso de la distribución hipergeométrica, la probabilidad de éxito cambia en cada ensayo, ya que la población se reduce a medida que se extraen elementos.

Además, en el limite donde la población N es grande y la probabilidad de éxito $p = \frac{K}{N}$ es constante, la distribución hipergeométrica se puede aproximar por una distribución binomial: $H(k|N, K, n) \simeq B(k|n, p = \frac{K}{N})$.

Relación entre Distribución de Poisson y Binomial La distribución de Poisson es una aproximación a la distribución binomial cuando el número de ensayos n es grande y la probabilidad de éxito p es pequeña, de tal forma que $np = \lambda$ se mantiene constante. En este caso, la función de probabilidad de la distribución binomial se puede aproximar por la función de probabilidad de la distribución de Poisson.

* mleize@df.uba.ar

2 Distribuciones continuas

Las distribuciones continuas se utilizan para modelar variables aleatorias que pueden tomar cualquier valor en un intervalo continuo. En las distribuciones continuas se utiliza la función de densidad de probabilidad (PDF) para describir la probabilidad de que una variable aleatoria X tome un valor en un intervalo específico. En una dimension, la PDF $f_X(t)$ se define como una función no negativa que integra a uno de la forma

$$\int_{-\infty}^{\infty} f_X(t)dt = 1\,. \tag{1}$$

La PDF se integra en un intervalo para obtener la probabilidad de que la variable aleatoria tome un valor en ese intervalo. La función de distribución acumulativa (CDF) se utiliza para describir la probabilidad de que una variable aleatoria tome un valor menor o igual a un valor específico

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt\,. \tag{2}$$

Es importante notar que la PDF no es una probabilidad, sino una densidad de probabilidad. La probabilidad de que una variable aleatoria tome un valor específico es cero

$$P(X = x) = \int_{x}^{x} f_X(t)dt = 0\,, \tag{3}$$

pero la probabilidad de que tome un valor en un intervalo específico es igual al área bajo la curva de la PDF en ese intervalo. Finalmente, notar que por este mismo motivo cuando hablamos de distribuciones continuas es análogo escribir

$$P(X \leq x) = P(X < x)\,. \tag{4}$$

La esperanza de una variable aleatoria continua se define como la integral de la variable

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx\,, \tag{5}$$

mientras que la varianza se define como la integral de la diferencia entre la variable y su esperanza al cuadrado

$$Var(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)dx\,. \tag{6}$$

3 Proceso de Poisson

El proceso de Poisson es un modelo estocástico que describe la ocurrencia de eventos en un intervalo de tiempo o espacio. Se caracteriza por la propiedad de que los eventos ocurren de forma independiente y a una tasa constante.

Distribución de Poisson La probabilidad de obtener k exitos en un proceso de Poisson está dada por la distribución de Poisson, que es una distribución discreta[1].

La función de probabilidad es $P(X = k) = \frac{\mu^k e^{-\mu}}{k!}$, donde μ es el número de eventos esperado en un intervalo dado y k es el número de eventos que ocurren en dicho intervalo. La media y la varianza de un proceso de Poisson son iguales y se denotan por μ. Si $\mu = \lambda \Delta t$, λ es la tasa media de ocurrencia de eventos en un intervalo de tiempo Δt.

Distribución Exponencial La distribución exponencial es una distribución continua que se utiliza para modelar el tiempo entre eventos en un proceso de Poisson. La función de densidad de probabilidad de la distribución exponencial es $f_X(t) = \lambda e^{-\lambda t}$, donde λ es la tasa de ocurrencia media de eventos.

[1] A veces van a encontrar que las funciones de densidad discretas (como las que vimos para la binomial y la hipergeometrica) se las define como Probability Mass Functions (PMF) y están dadas por $p_X(x) = P(X = x)$