Algunas Distribuciones de uso común en Probabilidad

Dra. Isabel Patricia Aguilar Juárez

Ensayo de Bernoulli

Sin lugar a dudas, el caso más sencillo de fenómeno aleatorio es aquel que tiene únicamente dos resultados posibles –satisfactorio o insatisfactorio, alto o bajo, lento o rápido, bueno o malo, etc. – que en general pueden denominarse

éxito (e) o ***fracaso (f)***

teniendo asociada, cada uno de ellos, una probabilidad de ocurrencia **p** y **q = 1 – p** respectivamente.

Este tipo de fenómeno recibe el nombre de

experimento o ensayo de Bernoulli.

Distribución de Bernoulli

Si denotamos con X a la v. a. que representa el número de éxitos que se presentan al realizar un ensayo de Bernoulli, se puede determinar inmediatamente que esta v. a. únicamente puede tomar los valores 0 y 1.

Nótese que la probabilidad de que no se observe ningún éxito, es exactamente la probabilidad de que el resultado sea un fracaso, es decir, 1 – p, y la probabilidad de obtener un éxito es p, esto es:

$$P(X = 0) = 1 - p \qquad \text{y} \qquad P(X = 1) = p$$

de donde la función de probabilidad asociada a la v. a. X definida anteriormente, es la función que tiene a p (la probabilidad de éxito), como único parámetro, y se conoce como

Distribución de Bernoulli.

Distribución de Bernoulli

Una forma común de escribir a la distribución de Bernoulli es la siguiente:

$$f_X(x) = \begin{cases} p^x(1-p)^{1-x} \ , & x = 0,1 \\ 0 \ , & en\ otro\ caso \end{cases}$$

El hecho de que una variable aleatoria X tenga la distribución de Bernoulli con parámetro p, se denota como $X \sim Ber(p)$

Las características numéricas de una v. a. que tenga esta distribución se pueden obtener, independientemente de la aplicación que se le de, desde luego, dichas características estarán en función del parámetro de la distribución. Así, es posible, partiendo de la definición, demostrar que

E(X) = p y Var(X) = p (1 – p).

Ensayo Binomial

Un **experimento o ensayo binomial** es una sucesión de ensayos de Bernoulli, todos independientes entre sí y con la misma probabilidad de éxito p .

Dentro de un ensayo binomial es posible plantearse varias preguntas tipo, que regularmente resultan de interés, como son:

- ¿Cuántos éxitos se observarán al realizar n repeticiones independientes de un ensayo de Bernoulli ?
- ¿Cuántas repeticiones del experimento se necesitarán para lograr obtener el primer éxito ?
- ¿Cuántas veces se tendrá que repetir el experimento, si se desea observar los primero r éxitos ?

Distribución Binomial

Llamemos X a la v. a. que representa el ***número de éxitos*** que se observan al realizar $\boldsymbol{n}$ ***ensayos*** independientes de Bernoulli, cada uno con la misma probabilidad de éxito $\boldsymbol{p}$, que todos los demás.

Observemos que, el rango de X son los enteros positivos entre 0 y n, inclusive.

En n ensayos de Bernoulli independientes, la forma de obtener exactamente $\boldsymbol{x}$ éxitos, en general no es única, puesto que si acomodamos en un arreglo los resultados obtenidos en cada ensayo, nos daremos cuenta de que existen exactamente ${}_nC_x = \binom{n}{x}$ arreglos diferentes que contengan el mismo número de éxitos. Lo anterior significa que la probabilidad buscada es la suma de las probabilidades de cada uno de los arreglos con la propiedad deseada.

Distribución Binomial

Por lo tanto, la distribución de probabilidad de X está dada por la siguiente función de probabilidad.

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \ldots, n \\ 0, & en\ otro\ caso \end{cases}$$

conocida como ***Distribución Binomial*** con parámetros $\boldsymbol{n}$ y $\boldsymbol{p}$ **.**

Cuando una variable aleatoria X tiene a esta función de probabilidad como distribución, se denota como $X \sim Bin(n, p)$.

Además, se puede demostrar que

$$E(X) = np \qquad \text{y} \qquad Var\,(X) = np(1-p)$$

Ejemplo 1

Suponga que los motores de un avión operan en forma independiente y que fallan con una probabilidad igual a 0.4. Considerando que un avión realiza un vuelo con toda seguridad si por lo menos trabaja la mitad de sus motores, determine la probabilidad de que un avión tetramotor efectúe un vuelo sin problemas.

Resolución

Sea X la v. a. que representa el número de motores que fallan cuando el avión realiza un vuelo.

Se observa que X es de la forma número de éxitos en n ensayos. Por tanto, $X \sim Bin(n, p)$

Dado que el avión es tetramotor, $n = 4$ y, de acuerdo con la información en el enunciado, $p = 0.4$, ya que es la probabilidad de que el motor funciones adecuadamente.

Por lo tanto, la probabilidad de que realice un vuelo seguro es $P(X \leq 2)$.

Como X es discreta,

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= \sum_{x=0}^{2} \binom{n}{x} p^x (1-p)^{n-x}$$
$$= \sum_{x=0}^{2} \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}$$

= 0.8208

Por lo tanto la probabilidad de realizar un vuelo seguro es más alta para el avión bimotor que para el tetramotor, aunque la dife-rencia es bastante pequeña.

Distribución Geométrica

número de veces que se debe repetir el experi-mento es al menos 1, pero es posible que se requieran más repeticiones. Por lo tanto los valores que puede tomar la variable aleatoria son todos los enteros positivos. Esto es $R_X = \{ 1, 2, 3, 4, \dots \}$.

Es claro que si se requiere realizar x veces el experimento con el fin de observar el primer éxito, es porque en las primeras (x – 1) repeticiones el resultado fue un fracaso, y solamente en el x – ésimo, el resultado es un éxito. La probabilidad de este resultado es $p\,(1 - p\,)^x$. Por lo tanto la función de probabilidad de X es

$$f_X(x) = \begin{cases} p\,(1 - p\,)^{(x-1)} & , \quad x = 1, 2, 3, \dots \\ 0 & , \quad \text{en otro caso} \end{cases}$$

que recibe el nombre de **Distribución Geométrica** con pará-metro **p.** La forma de denotar el hecho de que la v.a. X tiene como función de probabilidad a la distribución geométrica con parámetro p es X ~ G(p).

Las principales características numéricas de X son

$E(X) = 1 / p$ y $\quad \mathrm{Var}(X) = \dfrac{1 - p}{p^2}$

Distribución Binomial Negativa

Consideremos ahora la variable aleatoria X que representa el **número de veces** que se debe repetir el experimento, **con el fin de observar los primeros r éxitos**.

El rango de la variable aleatoria X es $R_X = \{ r, r+1, r+2, \ldots \}$.

Si se requieren x repeticiones del experimento para lograr observar los primeros r éxitos, esto significa que de los x ensayos **r** son éxitos y **x – r** son fracasos. Sin embargo, solamente el r-ésimo éxito está fijo en el último ensayo, los (r – 1) éxitos restantes se encontrarán en igual número de ensayos, pero existen exactamente formas distintas de obtener un resultado así. Por lo tanto, la función de probabilidad de X es

$$f_X(x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{(x-r)} & , \quad X = r, r+1, r+2 \ldots. \\ 0 & , \quad \text{en otro caso} \end{cases}$$

que recibe el nombre de **Distribución Geométrica** con pará-metro **p.**
La forma de denotar el hecho de que la v.a. X tiene como función de probabilidad a la distribución geométrica con parámetro p es X ~ G(p).

Las principales características numéricas de X son

$$E(X) = r / p \text{ y} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

Distribución de Poisson

El **Proceso de Poisson** consiste en la ocurrencia aleato-ria de eventos a lo largo de un intervalo continuo, de mane-ra que se reúnen las siguientes características:

Estacionaridad: La probabilidad de que ocurra un evento en un intervalo de longitud pequeña t, es λt, con λ constan-te.

Unicidad: La probabilidad de que ocurra más de un evento en un intervalo muy pequeño de longitud h es despreciable comparada con la probabilidad de que ocurra solo uno.

Independencia: La ocurrencia de un evento en un interva-lo dado, no depende de la ocurrencia en intervalos anterio-res, ni posteriores.

Sea X la v.a. que representa el **número** total **de eventos** que ocurren **en un intervalo de longitud t**.

Como se puede ver, X es una v. a. discreta cuyo rango es $R_X = \{ 0, 1, 2, 3, \ldots \}$, y su función de probabilidad es

$$f_X(x) = \begin{cases} \dfrac{(\lambda t)^x}{x!} e^{-\lambda t} & ; \quad x = 0, 1, 2, \ldots.. \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

Además, $E(X) = \lambda t$ en tanto que $Var(X) = \lambda t$

Ejemplo 2

Suponga que las moléculas de un gas raro se encuentran a razón promedio de tres por pie cúbico de aire. Si se su-pone que las moléculas están distribuidas independiente-mente y al azar en el aire, ¿cuál es la probabilidad de que, en una muestra de un pie cúbico de aire, se encuentre como máximo una molécula de dicho gas?

Resolución

Sea X la v. a. que representa el número de moléculas del gas raro, que se encuentran en un pie cúbico de aire.

X tiene distribución de Poisson con parámetro $\lambda = 3$, lo cual se denota como $X \sim P\,(\lambda t)$ con $\lambda = 3$ y $t = 1$.

La probabilidad que se desea calcular es $P\,(X \leq 1)$, lo cual se obtiene como sigue:

$$P\,(X \leq 1) = P\,(X = 0) + P\,(X = 1)$$
$$= \frac{\lambda t^{0} e^{-\lambda t}}{0\,!} + \frac{\lambda t^{1} e^{-\lambda t}}{1\,!} = 0.199$$

Se puede demostrar que en general la distribución de Poisson resulta una buena aproximación a la distribución binomial, si n es grande y p pequeña. Una regla práctica admisible es utilizar dicha aproximación si $n \geq 20$ y $p \leq 0.05$.

Distribución Exponencial

Definamos a la v, a. T como el tiempo que transcurre entre dos ocurrencias sucesivas de un evento, en un proceso de Poisson.

Nótese que la variable aleatoria puede adoptar cualquier valor real positivo, es decir, $R_T = \mathbb{R}^+$.

Una forma sencilla para conocer su distribución de probabilidad, consiste en relacionarla con la v. a. X con distribución de Poisson analizada antes. Al hacerlo, construyendo la función de distribución (acumulativa) de T y derivando con respecto a t, se obtiene la función de densidad de T, como

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & , \quad t > 0 \\ 0 & , \quad en\ otro\ caso \end{cases}$$

Conocida como **Distribución Exponencial** con parámetro λ. El hecho de que una v. a. T tenga esta distribución se indica como $T \sim \exp(\lambda)$.

Además si una v. a. T tiene este tipo de distribución, se puede calcular que

$$E(T) = 1/\lambda \qquad y \qquad Var(T) = 1/\lambda^2$$

Distribución Normal

La distribución más importante en el campo de la estadística, es la distribución Normal. Su gráfica es una curva en forma de campana, que describe la distribución de muchos conjuntos de datos que ocurren en la naturaleza, la industria y la investigación.

La expresión analítica más simple cuya gráfica tiene la forma acampanada, es la Gaussiana dada por

$$y = e^{-\frac{x^2}{2}}$$

Para que la campana Gaussiana se adapte a la población, es necesario introducir tres parámetros:

a: parámetro de localización que centre la gráfica en la media de la distribución de la población,
b: parámetro de contracción o dilatación que determine el ancho de la curva,
c: parámetro de normalización para que el área bajo la curva sea unitaria.

De manera que

$$f_X(x) = ce^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} \quad , \qquad x \in \mathbb{R}$$

en donde es claro que $\boldsymbol{a}$ debe ser la media y b la desviación estándar de la población, esto es, $\boldsymbol{a} = \mu$ y $\boldsymbol{b} = \sigma^2$.

Distribución Normal

Para obtener el valor de c, se utiliza el hecho de que la curva debe integrar uno

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} c e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx$$

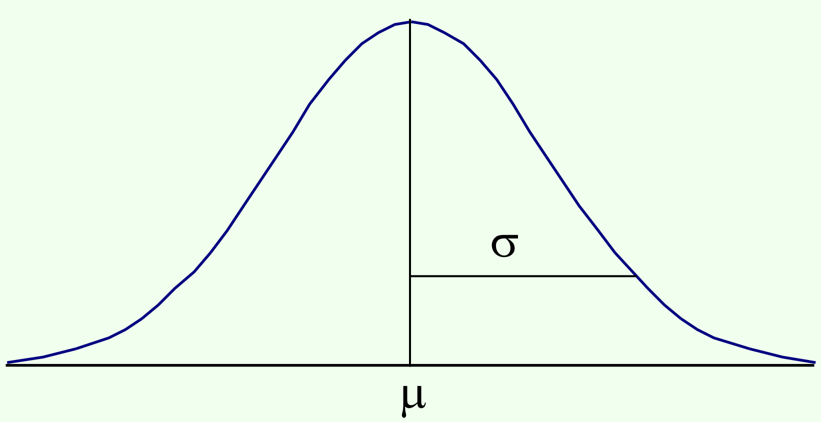
$$c = \frac{1}{b\sqrt{2\pi}}$$

Por lo que la expresión de la función de densidad Normal es :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{si} \quad -\infty < X < \infty$$

Esta función se conoce como **Distribución Normal** con media μ y varianza σ_2 y tiene la característica de ser simé-trica con respecto a la media. El hecho de que una v.a. tenga a esta función como distribución se denota como

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



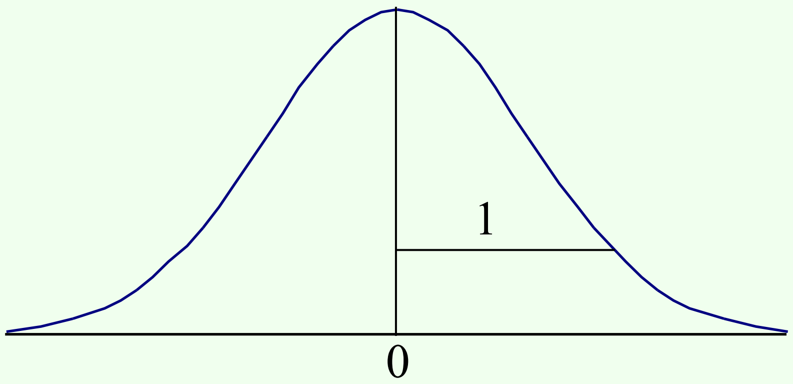
Distribución Normal Estándar

Se llama distribución normal estándar a aquella distribución normal, cuyos parámetros son $\mu = 0$ y $\sigma^2 = 1$. Esta distribución es tan importante, que la v.a. que tiene esta distribución, recibe el nombre especial de Z. Por lo tanto, la función de densidad es

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad -\infty < X < \infty$$

y se denota como

$$Z \sim N(0, 1)$$



Procedimiento de Estandarización

Si $X \sim N(\mu, \sigma^2)$, entonces la v.a.

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

Lo que significa que:

$P(x_1 < X < x_2) = P(z_1 < Z < z_2)$, en donde $z_1 = \dfrac{x_1 - \mu}{\sigma}$,

y $z_2 = \dfrac{x_2 - \mu}{\sigma}$