“Introducción a la Estadística, Probabilidad e Inferencia”

Maestría en Estadística Aplicada

Facultad de Ciencias Económicas y Estadística

UNR

Unidad 4 – Parte 1

- Variables aleatorias unidimensionales. Casos discreto y continuo.
- Funciones de probabilidad y de densidad. Función de distribución.
- Esperanza y variancia de una variable aleatoria.

Introducción

En la unidad anterior trabajamos sobre experimentos aleatorios, definiendo sus espacios muestrales y asignando probabilidad a eventos sobre tales espacios.

En muchas ocasiones, los resultados incluidos en el espacio muestral no son necesariamente los resultados de interés, en su lugar interesa *medir* algo sobre cada uno de los resultados, es decir, interesa ***asignar a cada resultado posible de Ω un número real que represente una característica de interés.***

Por ejemplo, sea un experimento en el que se selecciona 1 persona al azar de una cierta población. El espacio muestral estará compuesto por todos y cada uno de los individuos de la población. Sobre cada uno de ellos puede interesar medir alguna característica, peso, talla, edad, ingreso salarial, si fuma o no, etc.

Variables aleatorias

Definición: Sea E un experimento aleatorio y Ω su espacio muestral. Una función X que asigna a cada uno de los elementos $\omega \in \Omega$ un número real $X(\omega)$, se llama ***variable aleatoria***.

OBS: hay que tener en cuenta que no toda función que se conciba puede ser una variable aleatoria, deben cumplirse algunos requisitos. No obstante, dado que en la mayoría de las aplicaciones no se presentan inconvenientes en cuanto al cumplimiento de tales requisitos, no entraremos en detalle sobre ello ahora.

En algunos casos, los resultados del experimento constituyen directamente la característica numérica de interés, en cuyo caso, tomaremos como variable aleatoria a la función identidad: $X(\omega) = \omega$.

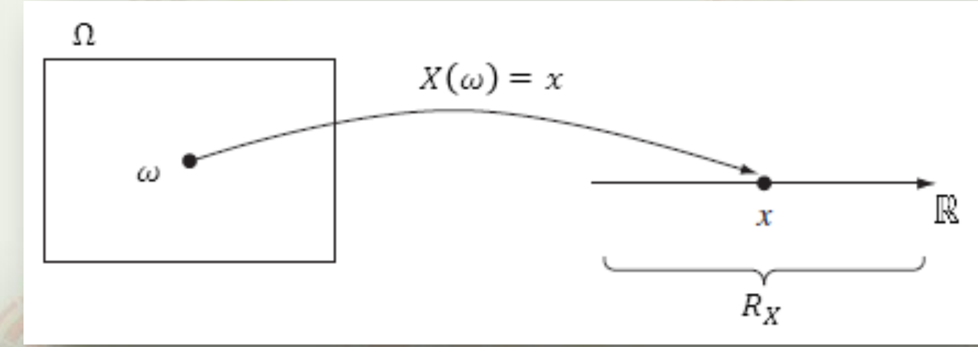
Variables aleatorias

Podemos pensar que las variables aleatorias "traducen" los resultados del espacio muestral a números reales, mediante la función:

$$X(\omega): \Omega \rightarrow \mathbb{R} |\, X(\omega) = x$$

Dado que las variables aleatorias se definen como funciones, será necesario que a cada $\omega \in \Omega$ le corresponda exactamente un valor $X(\omega)$.

El conjunto de todos los valores posibles de X se denomina recorrido de X y se simboliza usualmente R_X.



Notación:

- Variables aleatorias: X, Y, Z, …
- Sus posibles valores: x, y, z, …

Variables aleatorias - Sucesos

Así como establecimos que muchas veces es de interés algún suceso sobre los resultados de un experimento aleatorio, también puede definirse la noción de "suceso" con respecto a los resultados de la variable aleatoria X, y ellos estarán formados por subconjuntos del recorrido de X.

Definición: sea E un experimento aleatorio, Ω su espacio muestral, X una variable aleatoria en Ω y R_X su recorrido. Sea B un suceso respecto a R_X ($B \subset R_X$).
Si consideramos el conjunto $A = \{\omega \in \Omega | X(\omega) \in B\}$, decimos que A y B son ***sucesos equivalentes***.

Esto es, dos sucesos A y B son equivalentes siempre que ocurran juntos, aunque es importante señalar que cada uno está asociado a un espacio diferente.

Variables aleatorias - Sucesos

Ejemplo:

Sea E: lanzar una moneda dos veces al aire, y X la variable aleatoria que cuenta el número de caras obtenidas. Entonces:

$$\Omega = \{CC, XC, CX, XX\}$$
$$R_X = \{0, 1, 2\}$$

Sea el suceso $B_1 = \{1\}$, puesto que $X(\omega) = 1$ si y solo si $\omega = XC$ o $\omega = CX$, los sucesos $A_1 = \{CX, XC\}$ y $B_1 = \{1\}$ son equivalentes.

De la misma manera, $A_2 = \{CC\}$ y $B_2 = \{2\}$ así como $A_3 = \{XX\}$ y $B_3 = \{0\}$ son sucesos equivalentes.

¿Cómo asignar probabilidades a los sucesos en este nuevo espacio?

Variables aleatorias - Sucesos

<u>Definición:</u> sea B un suceso en el recorrido de X, R_X. La ***probabilidad del suceso B*** se define como la probabilidad del suceso $A \subset \Omega$ que es equivalente a B. Esto es:

$$P_X(B) = P(A) \quad \text{donde } A = \{\omega \in \Omega | X(\omega) \in B\}$$

es decir, A es el evento equivalente a B.

Esta definición permite asignar probabilidades a sucesos definidos en R_X en función de las probabilidades en Ω. Las probabilidades en R_X se denominan ***probabilidades inducidas*** por la variable aleatoria X.

<u>Ejemplo:</u> en el ejemplo del lanzamiento de dos monedas, suponiendo que las monedas son regulares, tenemos que:

$$P(CX) = P(XC) = \frac{1}{4} \quad \Rightarrow \quad P(A) = \frac{1}{2}$$

Luego, dado que los sucesos A y B son equivalentes, resulta:

$$P_X(B) = P(X = 1) = \frac{1}{2}$$

Variables aleatorias

La información contenida en la función de probabilidad P es completamente equivalente a la contenida en la función de probabilidad inducida P_X. Conocida P es posible obtener P_X, y viceversa.

De este modo, una vez que se han *inducido* las probabilidades de los sucesos en R_X, se puede ignorar el espacio muestral original Ω que dio lugar a esas probabilidades.

Cuando trabajamos con variables aleatorias podemos distinguir dos casos importantes: ***variables aleatorias discretas*** y ***variables aleatorias continuas***.

Variables aleatorias discretas

Definición: Sea X una variable aleatoria. Si el número de valores posibles de X, es decir, R_X es finito o infinito numerable, diremos que X es una ***variable aleatoria discreta***.

Esto significa que si X es una variables aleatoria discreta, sus posibles valores pueden *enumerarse*.

Ejemplo: una fuente radioactiva emite partículas α. Un contador observa la emisión de esas partículas durante un tiempo determinado, siendo de interés la variable aleatoria:

X: número de partículas observadas en un período de tiempo.

¿Cuáles son los posibles valores de X?

Podríamos pensar en $\mathrm{R_X} = \{0, 1,2, 3, \ldots\} = \mathbb{N}_0$

Variables aleatorias discretas

¿Cómo resumir o describir el comportamiento en probabilidad de una variable aleatoria discreta?

Si X es una variable aleatoria discreta, su recorrido R_X consta de un número a lo sumo numerable de valores $x_1, x_2, x_3, \dots$. Como ya hemos visto, cada uno de estos valores tiene asociado un número $P(X = x_i)$ llamado probabilidad de x_i.

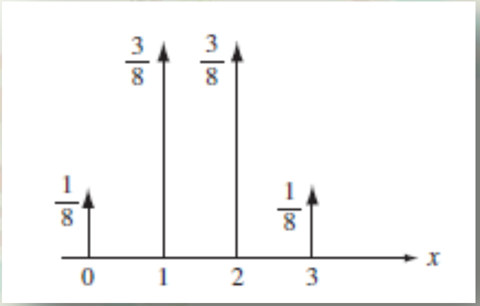
Definición: Si los números $p(x_i) = P(X = x_i)\ \forall x_i \in R_X$ satisfacen las condiciones:

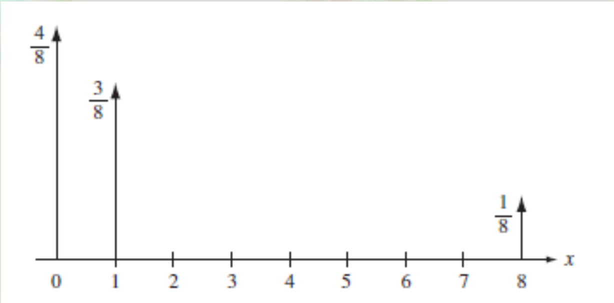
- $p(x_i) \geq 0\ \ \forall\, x_i \in R_X$
- $\sum_{i=1}^{\infty} p(x_i) = 1$

la función $p(x) = P(X = x)$ se llama ***función de probabilidad*** (puntual) de la variable aleatoria X, y lo notamos como: $\boldsymbol{X \sim p(x)}$

Variables aleatorias discretas

Representación gráfica:





Observaciones:

- Si la variable sólo toma un número finito de valores distintos, $x_1, x_2, \dots, x_N$, y todos ellos son igualmente probables resulta: $p(x_1) = p(x_2) = \cdots = p(x_N) = \frac{1}{N}$
- Si la variable toma un número infinito numerable de valores distintos, entonces ellos no pueden ser equiprobables.
- Si B es un evento asociado con la variable aleatoria X, es decir, $B = \{x_i,\ i \in I\} \subset R_X$, entonces:

$$P_X(B) = P(\omega \in \Omega | X(\omega) = x_i, \forall\, i \in I) = \sum_{i \in I} p(x_i)$$

Variables aleatorias discretas

Ejemplo:

Se realizan pruebas de un componente eléctrico de una máquina, cuya probabilidad de falla es ¼. Se prueban sucesivamente muchos componentes, hasta que se encuentra uno que falla (F). Sea la variable aleatoria X que cuenta el número de pruebas hasta encontrar el primer fallo.

$$\Omega = \{F, \bar{F}F, \bar{F}\bar{F}F, \bar{F}\bar{F}\bar{F}F, \bar{F}\bar{F}\bar{F}\bar{F}F, \dots\} \quad \text{y} \quad R_X = \{1, 2, 3, 4, \dots\}$$

***¿Como encontrar la función de probabilidad de X*?**

$X = k \Leftrightarrow$ los primeros $k - 1$ componentes no fallan y el k-ésimo falla.

$$P(X = k) = \left(\frac{3}{4}\right)^{k-1} \frac{1}{4}$$

Sea el suceso B_1: son necesarios más de *3* ensayos para encontrar uno que falla. ***¿Cuál es la probabilidad de $\boldsymbol{B_1}$?***

Si $\boldsymbol{B_2}$ es el suceso definido como *la prueba termina luego de un número par de ensayos*, ¿cuál es su probabilidad?

Variables aleatorias continuas

Supongamos ahora que la variable aleatoria X puede tomar todos los valores posibles, en un determinado intervalo $I \in \mathbb{R}$.

¿Qué sucede con las probabilidades puntuales $\boldsymbol{p(x_i)}$?

Puesto que los valores de X no son contables, no podemos hablar del i-ésimo valor x_i y por lo tanto $p(x_i)$ pierde sentido.
En estos casos, corresponde sustituir la función de probabilidad $p(x)$ por una función f definida para todos los valores posibles de X.

Definición: Diremos que X es una ***variable aleatoria continua*** si existe una función f llamada ***función de densidad de probabilidad (fdp)*** que satisface las siguientes condiciones:

- $f(x) \geq 0 \quad \forall\, x$
- $\int_{-\infty}^{\infty} f(x)dx = 1$

en cuyo caso lo notaremos como $\boldsymbol{X \sim f(x)}$.

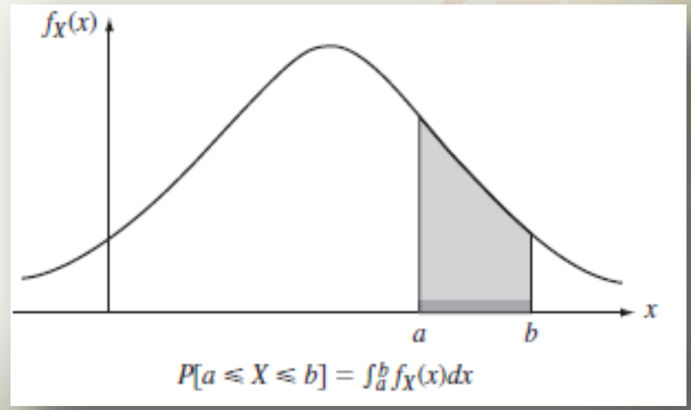
Variables aleatorias continuas

Observaciones:

- Las variables aleatorias continuas pueden tomar valores en todo el eje real o bien en ciertos intervalos de número reales. Si X sólo toma valores en un intervalo $[a, b]$, se establece $f(x) = 0$ $\forall x \notin [a, b]$, de modo que la fdp está definida para todos los valores reales de X.
- Para cualesquiera a y b tales que $[a, b] \subset I$, la probabilidad de que X tome valores en $[a, b]$ es:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

- Este valor de probabilidad representa el área bajo la curva definida por la fdp $f(x)$ entre los valores $x = a$ y $x = b$.



Variables aleatorias continuas

Observaciones:

- Una consecuencia de la definición dada es que para cualquier valor puntual de X, digamos x_0, tenemos: $P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$

- Como consecuencia de este hecho, si X es una variable aleatoria continua, las siguientes probabilidades son todas iguales:

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

- **El valor de $\boldsymbol{f(x)}$ NO ES UNA PROBABILIDAD.** Sólo cuando la fdp se integra entre dos límites produce una probabilidad. Entonces, ***¿qué representa $\boldsymbol{f(x)}$***?

Variables aleatorias continuas

Observaciones:

La pdf en un punto x_0 representa la ***"densidad"*** de probabilidad en dicho punto. Veamos: la probabilidad que X tome valores en un pequeño intervalo en la vecindad de x_0 es:

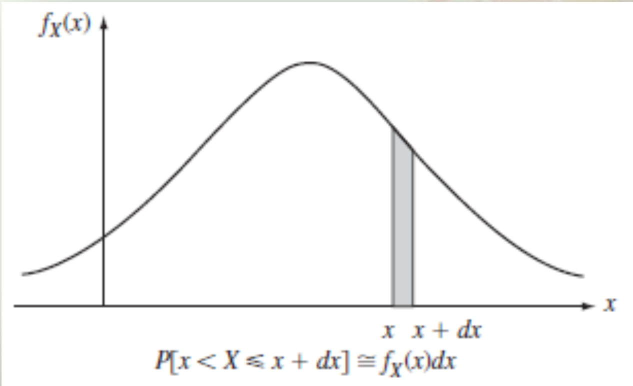
$$P(x_0 \leq X \leq x_0 + \Delta_x) = \int_{x_0}^{x_0+\Delta_x} f(x)dx = f(\xi)\Delta_x \, ; \qquad x \leq \xi \leq x + \Delta_x$$

Entonces, si Δ_x es pequeño, podríamos escribir:

$$P(x_0 \leq X \leq x_0 + \Delta_x) \approx f(x)\Delta_x$$

de modo que $f(x)$ representa el peso o la densidad de probabilidad que existe en el entorno del punto x.

Si $f(x_i) > f(x_j)$, entonces la probabilidad de un intervalo de amplitud Δ_x alrededor de x_i es mayor que la de un intervalo de la misma amplitud alrededor de x_j.



Variables aleatorias continuas

Ejemplo: Sea X una variable aleatoria continua con fdp dada por:

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & en\ otro\ caso \end{cases}$$

Claramente $f(x)$ es una función de densidad pues $f(x) \geq 0$ y se verifica la condición de cierre: $\int_0^1 2xdx = 1$.

Entonces podemos calcular la probabilidad de eventos en R_X, integrando esta fdp. Por ejemplo:

$$P\left(X < \frac{1}{2}\right) = \int_0^{1/2} 2xdx = \frac{1}{4}$$

O incluso, calcular probabilidades condicionales:

$$P\left(X < \frac{1}{2} \middle| \frac{1}{3} < X \leq \frac{2}{3}\right) = \frac{P\left(\frac{1}{3} < X < \frac{1}{2}\right)}{P\left(\frac{1}{3} < X \leq \frac{2}{3}\right)} = \frac{\int_{1/3}^{1/2} 2xdx}{\int_{1/3}^{2/3} 2xdx} = \frac{5/36}{1/3} = \frac{5}{12}$$

Función de distribución acumulada

Definición: Sea X una variable aleatoria, discreta o continua. La ***función de distribución acumulada (fda)*** de la variable aleatoria X se define, para cada $x \in \mathbb{R}$ como la probabilidad del evento $\{X \leq x\}$. Esto es:

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

En virtud de los axiomas de la probabilidad, la función de distribución acumulada goza de las siguientes ***propiedades***:

- $F_X(x)$ es una función no decreciente de x:

  si $x_1 < x_2$ entonces $F_X(x_1) \leq F_X(x_2)$.

- $F_X(x)$ es una función continua por derecha:

  dado $h > 0$, $F_X(b) = \lim_{h \to 0} F_X(b+h)$.

- $F_X(x)$ está acotada entre 0 y 1:

  $\lim_{x \to -\infty} F_X(x) = 0$ y $\lim_{x \to \infty} F_X(x) = 1$.

Función de distribución acumulada

Si X es una variable aleatoria discreta, entonces:

$$F_X(x) = \sum_{x_i \in R_X | x_i \leq x} p(x_i)$$

Esto significa que si X es variable aleatoria discreta, su función de distribución se formará de trazos horizontales, con saltos en los puntos x_i, formando una función de tipo escalonada, continua a derecha y cuyos saltos serán de magnitud $p(x_i)$.

Ejemplo: sea el experimento aleatorio consistente en lanzar al aire una moneda tres veces.
Sea X la variable aleatoria que cuenta el número de caras en las tres tiradas. Entonces: $R_X = \{0, 1, 2, 3\}$ con probabilidades 1/8, 3/8, 3/8 y 1/8 respectivamente (comprobarlo!).

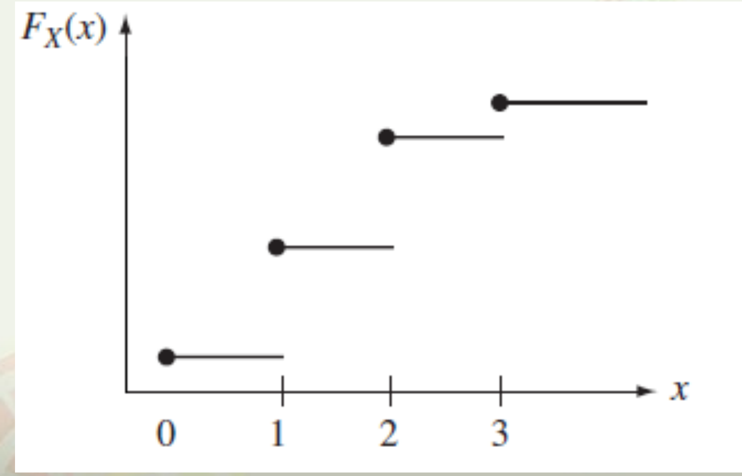
¿Cómo hallamos la función de distribución?

Función de distribución acumulada

Dado que X es una variable aleatoria discreta: $F_X(x) = \sum_{x_i \leq x} p(x_i)$

Entonces:

- Si $x < 0$: $F_X(x) = 0$, pues X no toma valores menores a cero.
- Si $x = 0$: $F_X(0) = P(X \leq 0) = P(X < 0) + P(X = 0) = p(0) = 1/8$,
- Si $0 < x < 1$: $F_X(x) = P(X \leq x) = P(X \leq 0) + P(0 < X \leq x) = p = 1/8$, pues X no toma valores en el intervalo $(0, 1)$.
- Si $x = 1$: $F_X(1) = P(X \leq 1) = P(X < 1) + P(X = 1) = p(0) + p(1) = 4/8$
- …



Función de distribución acumulada

Si X es una variable aleatoria continua, entonces:

$$F_X(x) = \int_{-\infty}^{x} f(x)dx$$

Si X es variable aleatoria continua, su función de distribución será una función continua para todo valor de x, su gráfica no presentará cortes ni saltos.

Ejemplo: Sea X la variable aleatoria que representa el tiempo que tarda un sistema de telecomunicaciones en transferir una llamada. Por su naturaleza, se trata de una variable aleatoria continua, cuya función de densidad supongamos que es: $f(x) = \lambda e^{-\lambda x} \quad \forall \ x \geq 0$.

Modelo exponencial

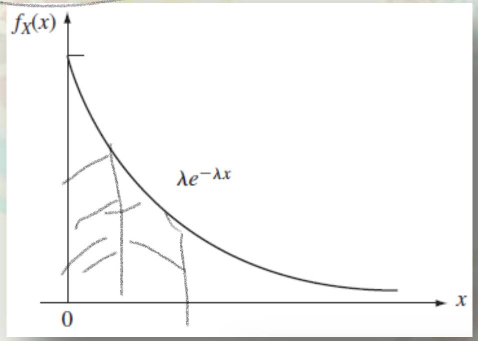
Función de distribución acumulada

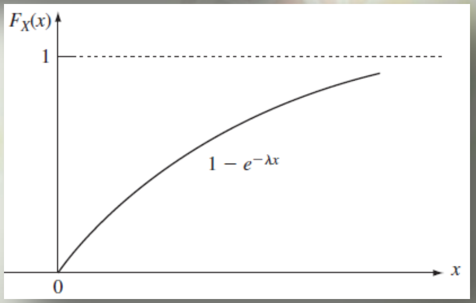
Dado que X es una variable aleatoria continua: $F_X(x) = \int_{-\infty}^{x} f(x)dx$

Entonces:

- Si $x < 0$: $f(x) = 0$, luego $F_X(x) = 0$.
- Si $x \geq 0$: $F_X(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{x} \lambda e^{-\lambda x}dx = 1 - e^{-\lambda x}$

Por lo tanto: $F_X(x) = \begin{cases} 0 & si\ x < 0 \\ 1 - e^{-\lambda x} & si\ x \geq 0 \end{cases}$





Supongamos que $\lambda = 1$, ¿cuál es la probabilidad de que el sistema demore a lo sumo 1 minuto en transferir la llamada?

$\boldsymbol{P(X \leq 1) = F(1) = 1 - e^{-1} = 0.63}$

¿y de que demore más de 2 minutos?...

Función de distribución acumulada

Hemos visto que la función de distribución acumulada puede obtenerse a partir de la función de probabilidad en el caso discreto y de la función de densidad en el caso continuo.

También es posible hacer el camino inverso, dada una función de distribución es posible obtener la función de probabilidad o la función de densidad según corresponda. En efecto:

- Si F es la fda de una variable aleatoria **continua** con densidad f, entonces: $\boldsymbol{f(x) = \frac{d}{dx}F(x)}$ para todo valor de x en que F sea diferenciable.

- Si F es la fda de una variable aleatoria **discreta** que toma los valores $x_1, x_2, \ldots$ y tales valores están ordenados de modo que $x_1 < x_2 < \cdots$, entonces:

$$\boldsymbol{p(x_i)} = P(X = x_i) = P(X \le x_i) - P(X \le x_{i-1}) = \boldsymbol{F(x_i) - F(x_{i-1})}$$

Función de distribución acumulada

Vimos que es posible calcular probabilidades de eventos en R_X a partir de la función de distribución acumulada. En este sentido pueden resultar de utilidad las siguientes reglas:

- $P(a < X \leq b) = F_X(b) - F_X(a)$
- $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$
- $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$
- $P(a < X < b) = F_X(b^-) - F_X(a)$

- $P(X = b) = F_X(b) - F_X(b^-)$

- $P(X > b) = 1 - F_X(b)$
- $P(X \geq b) = 1 - F_X(b^-)$

Donde la notación $F_X(a^-)$ representa el límite de $F(x)$ cuando x tiende a a por izquierda: $F_X(a^-) = \lim_{x \to a^-} F_X(x)$.

OBS: notar que si X es variable aleatoria continua, las 4 primeras probabilidades listadas serán iguales entre sí, la quinta dará cero y las dos últimas también coincidirán entre sí.

Esperanza de una variable aleatoria

Con cada distribución de probabilidad podemos asociar ciertos parámetros que ayudan a caracterizar dicha distribución y resumen la información contenida en la misma.

Ejemplo:

Un fabricante de ciertos artículos sabe que las condiciones de operación actuales producen alrededor de un 4% de artículos defectuosos.
Por cada artículo defectuoso, el fabricante pierde \$10, mientras que la utilidad de los productos conformes es de \$50.
Si X es la variable que representa la utilidad neta por artículo, el fabricante desea saber cual es el valor de utilidad que puede esperar en su producción a largo plazo.

Esperanza de una variable aleatoria

Puesto que el fabricante pierde \$10 alrededor del 4% de las veces y gana \$50 en el 96% restante de las veces, parece razonable calcular:

$$\mathbf{-10 \cdot 0,04 + 50 \cdot 0,96 = 47,60}$$

Concluyendo que, a largo plazo, con esa tasa de producción de artículos defectuosos y con esos costos asociados, se espera tener una utilidad de \$47,60 por artículo.

El cálculo realizado corresponde a la ***esperanza*** o ***valor esperado*** de la variable aleatoria X, notada como $E(X)$.

Este valor representa el "centro de gravedad" de la distribución, corresponde al valor de la variable aleatoria que , en promedio, se espera observar a largo plazo.

Esperanza de una variable aleatoria

CASO DISCRETO:

Definición: sea X una variable aleatoria discreta con valores posibles $x_1, x_2, \dots, x_n, \dots$. Sea $p(x)$ la función de probabilidad de la variable aleatoria. La ***esperanza o valor esperado*** de X, $E(X)$, se define como:

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i)$$

si la serie del lado derecho converge absolutamente.

Observaciones:

- Si X toma solo un número finito de valores entonces la expresión anterior se reduce a una suma ordinaria.
- Ésta se puede interpretar como un promedio ponderado de los valores posibles de X, con pesos dados por las probabilidades puntuales.
- Si los resultados son equiprobables, la esperanza de X resulta el promedio aritmético ordinario de los valores posibles de la variable.

Esperanza de una variable aleatoria

CASO CONTINUO:

Definición: sea X una variable aleatoria continua con fdp $f(x)$. La ***esperanza o valor esperado*** de X, $E(X)$, se define como:

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

si la integral impropia converge absolutamente.

Ejemplo: recordemos el ejemplo sobre el tiempo de transferencia de las llamadas por el sistema de telecomunicaciones. Se tenía $X \sim f(x) = \lambda e^{-\lambda x}$ con $\lambda = 1$, $\forall\, x \geq 0$. Entonces,

$$E(X) = \int_0^{\infty} x\, e^{-x} dx = -x\, e^{-x}\Big|_0^{\infty} + \int_0^{\infty} e^{-x} dx = 1$$

Es decir, es de esperar que en promedio el sistema demore un minuto para transferir las llamadas.

Esperanza de una función de una v.a.

Sea $\boldsymbol{X}$ una variable aleatoria y sea $\boldsymbol{Y = g(X)}$.
Entonces (en general) ***Y también es una variable aleatoria***, y como tal tendrá su propia distribución de probabilidad.

Supongamos que el interés es evaluar la esperanza de esta nueva variable aleatoria Y. El camino natural es considerar la aplicación de la definición de esperanza a la variable Y haciendo uso de su distribución (desconocida, debería deducirse).

Sin embargo, dado que X e Y están funcionalmente relacionadas, resulta natural preguntarse si ***podremos obtener $E(Y)$ sólo a partir del conocimiento de la distribución de probabilidades de X***, sin necesidad de derivar la distribución de probabilidad de Y. La respuesta es afirmativa.

Esperanza de una función de una v.a.

Teorema: Sea X una variable aleatoria y sea $Y = g(X)$.

- Si X es una variable aleatoria discreta con $p(x_i) = P(X = x_i)$, entonces:

$$E(Y) = E\left(g(X)\right) = \sum_{i=1}^{\infty} g(x_i)p(x_i)$$

- Si X es una variable aleatoria continua con fdp $f(x)$, entonces:

$$E(Y) = E\left(g(X)\right) = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

Esperanza de una función de una v.a.

<u>Ejemplo:</u>

Sea V la velocidad del viento (km/h) con fdp $f(v) = \frac{1}{10}\ \ \forall v \in [0, 10]$.
La presión W (lb/pie^2) sobre la superficie del ala de un aeroplano está dada por la relación $W = 0{,}003V^2$.
Procediendo según el teorema, podemos hallar la esperanza de W como:

$$E(W) = E(0{,}003V^2) = \int_{-\infty}^{\infty} 0{,}003v^2 \frac{1}{10}\, dv = 0{,}1$$

Es decir, la presión a la que se espera que en promedio estén expuestos los aviones como resultado del viento es de 0,1 lb/pie^2.

Propiedades de la Esperanza

Supongamos que X es una variable aleatoria para la cual existe $E(X)$.

Propiedad 1: Esperanza de una función lineal.
Si $Y = aX + b$, con a y b constantes finitas, entonces $E(Y) = aE(X) + b$

En efecto, supongamos que X es variable aleatoria continua, entonces, por definición:

$$E(aX+b) = \int_{-\infty}^{\infty}(ax+b)f(x)dx = a\int_{-\infty}^{\infty}x\,f(x)dx + b\int_{-\infty}^{\infty}f(x)dx =$$
$$= aE(X) + b$$

De manera análoga se prueba para el caso continuo.

Propiedades de la Esperanza

Propiedad 2: Esperanza de una constante.
Si $X = c$, con c una constante finita, entonces $E(X) = c$.

Propiedad 3: Esperanza de una suma de variables aleatorias.
Sean X e Y dos variables aleatorias con esperanza finita. Entonces, $E(X+Y) = E(X) + E(Y)$.

Las propiedades 1 y 3 pueden combinarse y extenderse a la suma de más de dos términos, resultando la siguiente propiedad general:

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$$

Esta propiedad significa que la esperanza de una función lineal de variables aleatorias es igual a la misma función lineal de las esperanzas individuales. ***Esto no es cierto si la función no es lineal.***

Variancia de una variable aleatoria

El valor esperado de una variable aleatoria es uno de los parámetros mas importantes y utilizados a la hora de describir o caracterizar una distribución de probabilidad. Sin embargo, la información que provee no es del todo completa.
Si la esperanza de una variable aleatoria toma un valor, digamos k, eso significa que si consideramos un gran número de observaciones de la v.a., el promedio de tales valores se aproximará a k.

<u>Ejemplo:</u> Sea X una v.a. que representa la duración en horas de una lamparita y supongamos que se sabe que $E(X) = 1000$.

¿Alcanza esta información para caracterizar la distribución del tiempo de duración de las lamparitas?

Necesitamos una medida adicional que describa la ***dispersión de la variable aleatoria***.

Variancia de una variable aleatoria

<u>Definición:</u> sea X una variable aleatoria con esperanza finita. La ***variancia de X***, denotada por $V(X)$ se define como:

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

El ***desvío estándar*** de X se define como la raíz cuadrada no negativa de la variancia.

<u>Observaciones:</u>

- La variancia está expresada en unidades al cuadrado, mientras que el desvío estándar está expresado en las mismas unidades que la v.a.

- El cálculo de la variancia de una v.a. se simplifica bastante teniendo en cuenta la siguiente ***fórmula de trabajo***:

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

Variancia de una variable aleatoria

Ejemplo:

Caso discreto. Consideremos el ejemplo del fabricante de artículos que desea estimar la utilidad esperada.

Teníamos $E(X) = 47{,}60$, y además podemos calcular:

$E(X^2) = (-10)^2 \cdot 0{,}05 + 50^2 \cdot 0{,}95 = 2380.$

Luego:

$$V(X) = E(X^2) - [E(X)]^2 = 2380 - 47{,}60^2 = 114{,}24$$
$$DE(X) = \sqrt{114{,}24} = 10{,}69.$$

De manera similar puede aplicarse con variables aleatorias continuas, reemplazando las sumatorias por integrales.

Propiedades de la variancia de una v.a.

La función variancia de una variable aleatoria goza de ciertas propiedades:

Propiedad 1: Variancia del producto por una constante.
Si $Y = cX$, con c una constante finita, entonces $V(Y) = c^2V(X)$.

En efecto, a partir de la fórmula de trabajo tenemos:

$$V(cX) = E[(cX)^2] - [E(cX)]^2 = E(c^2X^2) - [cE(X)]^2 =$$
$$= \mathrm{c}^2\ [\mathrm{E}(\mathrm{X}^2) - [E(X)]^2] = \mathrm{c}^2\mathrm{V}(\mathrm{X})$$

Propiedad 2: Variancia de la suma con una constante.
Si $Y = X + c$, con c una constante finita, entonces $V(X + c) = V(X)$.

Estas propiedades indican que la variancia no se ve afectada por transformaciones de tipo traslación , pero sí por los cambios de escala.

Propiedades de la variancia de una v.a.

Propiedad 3: Variancia de una función lineal.
Combinado las propiedades 1 y 2 resulta que si $Y = cX + b$, con c y b constantes finitas, entonces $V(Y) = c^2V(X)$.

Propiedad 4: Variancia de la suma de variables aleatorias.
Para esta propiedad necesitaríamos definir distribución conjunta de variables aleatorias (lo veremos en la parte 3 de esta unidad) e independencia de variables aleatorias. No obstante adelantamos que, a diferencia de lo que sucede con la esperanza, la variancia de una suma de variables aleatorias cualesquiera NO es la suma de las variancias individuales. Esto es, la **variancia no es *aditiva***.

Sólo en caso que las variables sean independientes resulta:

$$V(X+Y) = V(X) + V(Y)$$

Momentos de una variable aleatoria

La esperanza y la variancia son los dos parámetros mas utilizados para resumir la distribución de probabilidad de una variable aleatoria. No obstante, existen otros parámetros utilizados en algunas situaciones.
El punto a notar es que tales parámetros involucran el valor esperado de potencias de mayor orden de la v.a. X. Estos valores esperados reciben el nombre de ***momentos de la variable aleatoria X.***

Definición:

- El ***k-ésimo momento*** de la v.a. X se define como: $\boldsymbol{\mu'_k = E(X^k)}$
- El ***k-esimo momento centrado*** de la v.a. X se define como: $\boldsymbol{\mu_k = E\left[\left(X - E(X)\right)^k\right]}$
- **El k-ésimo momento reducido** de la v.a. X se define como: $q_k = E\left[\left(\frac{X-E(X)}{\sqrt{V(X)}}\right)^k\right]$

Momentos de una variable aleatoria

Ejemplos:

- La **esperanza** de una variable aleatoria es el momento ordinario de orden 1: $\mu_1' = E(X)$
- Para cualquier variable aleatoria el primer momento centrado será igual a cero pues $\mu_1 = E[x - E(X)] = \mathrm{E(X)} - \mathrm{E(X)} = 0$.
- La **variancia** de una variable aleatoria es el momento centrado de orden 2: $V(X) = \mu_2 = E[x - E(X)]^2$
- El **coeficiente de asimetría** (indicador que establece el grado de asimetría que tiene una distribución) se corresponde con el momento reducido de orden 3: $q_3 = E\left[\left(\frac{X-E(X)}{\sqrt{V(X)}}\right)^3\right]$.

ANEXO - Cotas para la probabilidad

Hemos visto cómo, a partir de la distribución de probabilidad de una variable aleatoria, podemos calcular esperanza y variancia.

Sin embargo, si para una variable aleatoria sólo conocemos los valores de estos parámetros, no será posible con ellos reconstruir la distribución de probabilidad, pero sí podemos dar cotas para la probabilidad de ciertos tipos de eventos. Estos resultados están enmarcados en la conocida Desigualdad de Chebyshev.

Sea X una v.a. con esperanza finita y sea c un número real cualquiera. Entonces, si $E(X-c)^2$ es finita y ε es un número positivo cualquiera, tenemos:

$$P[|X-c| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} E(X-c)^2$$

ANEXO - Cotas para la probabilidad

Algunas formas equivalentes de la desigualdad:

- Considerando el suceso complementario:

$$P[|X - c| < \varepsilon] \geq 1 - \frac{1}{\varepsilon^2} E(X - c)^2$$

- Si se elige $c = E(X)$, entonces:

$$P[|X - E(X)| \geq \varepsilon] \leq \frac{V(X)}{\varepsilon^2}$$

- Si se elige $c = E(X)$ y $\varepsilon = k\sigma$, con $\sigma = \sqrt{V(X)}$ entonces:

$$P[|X - E(X)| \geq k\sigma] \leq \frac{1}{k^2}$$

Bibliografía

- Agresti A., Franklin C. (2009) The art and science of learning from data. 2º Ed. New Jersey. Pearson Prentice Hall.
- DeGroot M, Schervish M. (2012). Probability and Statistics. 4° Ed. Pearson Education Inc.
- Grimmett G.R., Stirzaker, D.R. (2001). Probability and Random Processes. 3º Ed. Oxford University Press.
- Meyer P. (1998). Probabilidad y Aplicaciones Estadísticas. Edición revisada. Addison-Wesley Iberoamericana.
- Ross S. (2010). A First Course in Probability. 8º Ed. Pearson Prentice Hall.
- Casella G., Berger R. (2002). Statistical Inference. 2° Ed. Duxbury Press.