Libro 2: Variables aleatorias discretas

Sitio:	Agencia de Habilidades para el Futuro	Imprimido por:	Eduardo Moreno
Curso:	Estadística y probabilidades para el desarrollo del soft 1º D	Día:	lunes, 7 de abril de 2025, 02:00
Libro:	Libro 2: Variables aleatorias discretas		

Tabla de contenidos

¿Por qué vamos a estudiar variables aleatorias?

En esta posta estudiaremos las **variables aleatorias**, haciendo énfasis en el caso **discreto**. Si bien lo explicaremos con más detalle luego, por ahora podemos pensar que la variable aleatoria permite asignar un único número a cada posible resultado de un experimento aleatorio.

Pero, ¿por qué las necesitaríamos? Vamos a ver que las variables aleatorias son una gran herramienta para modelar en el contexto de la probabilidad. ¿Por qué? Porque podremos representar y trabajar con conjuntos de resultados de manera más compacta (¡y abstracta!). Por ejemplo, en lugar de describir cada resultado individualmente, podremos asignar valores numéricos a esos resultados y estudiar las propiedades y comportamientos asociados a esos valores. ¡Y generalizar!

Además, en muchos escenarios del mundo real, los resultados no son equiprobables y pueden tomar valores distintos con diferentes probabilidades. Las variables aleatorias nos permiten capturar esta variabilidad y modelarla de manera más realista.

> **Por ejemplo** | Al lanzar una moneda cargada, cada cara de la moneda no tiene la misma probabilidad de aparecer, y una variable aleatoria nos permitirá representar esa variación. **De hecho, ese será un ejemplo que modelaremos en Python en la semana 5 del trabajo con esta posta.**

También veremos que las variables aleatorias nos facilitarán el cálculo de probabilidades asociadas a eventos particulares. Por ejemplo, en lugar de enumerar todos los resultados posibles de un experimento, podremos utilizar la función de probabilidad de la variable aleatoria para determinar la probabilidad de que ocurra un evento específico. Esto simplifica los cálculos y nos proporcionará una forma más eficiente de analizar situaciones en donde incide el azar.

Y todo esto, dará el puntapié inicial para el estudio de las distribuciones de probabilidad. Estas distribuciones serán fundamentales en estadística para comprender la variabilidad de los datos y elaborar conjeturas sobre poblaciones más grandes, a las que típicamente no tenemos acceso.

¡Pero no te preocupes! Iremos paso a paso. Por lo pronto, en esta primera introducción, te invitamos a ver un ejemplo muy sencillo de cómo definir una variable aleatoria a partir de lo que ya estudiamos en la posta 1.

¿Qué son las v ariables aleatorias?

Antes de enunciar formalmente la definición de variables aleatorias, les proponemos que **pensemos juntos el siguiente ejemplo.**

Ejemplo comentado

Supongamos que se prueban tres componentes electrónicos. En ese caso, puede ocurrir que cada componente sea "defectuoso" o "no defectuoso".

Según lo trabajado en la semana 1, en el libro "Introducción a la teoría de la probabilidad", el espacio muestral equiprobable que ofrece una descripción detallada de cada posible resultado, se puede escribir de la siguiente manera:

$$\Omega = \{NNN, NND, NDN, NDD, DNN, DND, DDN, DDD\}$$

donde N denota "no defectuoso" y D "defectuoso".

Analicemos el **número de componentes defectuosos** que se presentan en cada uno de los posibles resultados.

Por ejemplo, en el elemento NNN la cantidad de defectuosos es cero, pero en el NND es de uno. A esta información la podemos representar en una tabla como se muestra a continuación:

Elementos del espacio muestral	Cantidad de defectuosos
NNN	0
NND	1
NDN	1
NDD	2
DNN	1
DND	2
DDN	2
DDD	3

De esta forma, a cada punto en el espacio muestral se le asigna un valor numérico de 0, 1, 2 o 3. Estos valores son, por supuesto, cantidades determinadas por el resultado del experimento aleatorio.

¿Esta tabla no les resulta familiar? Es el mismo tipo de tabla que usábamos en Elementos de Análisis Matemático cuando estudiamos las **funciones**. Observen que las condiciones de unicidad y existencia se cumplen.

Entonces, la variable "número de artículos defectuosos" no es otra cosa que una función y recibe el nombre de **variable aleatoria**.

Definición formal de variable aleatoria

[recuadro] Una variable aleatoria es una función que asocia un número real a cada elemento del espacio muestra de un experimento aleatorio.

Notación

Utilizaremos una letra mayúscula, digamos X, para denotar una variable aleatoria, y su correspondiente letra minúscula, x en este caso, para referirnos a los posibles valores que toma esa variable. Es decir, mientras X denota a la variable aleatoria, x denota a un número que la variable pueda tomar.

Volviendo al ejemplo de los componentes electrónicos...

Observamos que la variable aleatoria X toma el valor 2 para todos los elementos en el subconjunto

$$E = \{DDN, DND, NDD\}$$

del espacio muestral Ω. Esto es, cada valor posible de X representa un evento que es un subconjunto del espacio muestral para el experimento dado.

El caso discreto

Vamos a analizar los siguientes ejemplos que nos permitirán comprender la definición de un tipo particular de variables aleatorias: las discretas.

Ejemplo 1

De una urna que contiene 5 pelotas rojas y 2 negras, se sacan dos pelotas de manera sucesiva, sin reemplazo.

Ejemplo comentado

Definimos la variable aleatoria Y como el número de pelotas rojas extraídas. Entonces, los posibles resultados junto con los valores y de la variable aleatoria Y, se muestran en la siguiente tabla:

Espacio muestral	y
RR	2\
RN	1
NR	1
NR	0

Por lo tanto, el rango de esta función es el conjunto $R_Y = \{0, 1, 2\}$, lo cual tiene sentido ¿no? Porque puede ser que ninguna pelota que salga sea roja, que sea una de las dos o las dos.

Ejemplo 2

Supongamos que un plan de muestreo implica obtener una muestra de artículos de un proceso hasta que se encuentre uno defectuoso. La evaluación del proceso dependerá de cuántos artículos consecutivos se observen.

Ejemplo comentado

En este caso, sea X una variable aleatoria que se define como el número de artículos observados antes de que salga uno defectuoso.

Si N representa un artículo "no defectuoso" y D uno "defectuoso", los espacios muestrales son:

- $\Omega_1 = \{D\}$, dado que $X = 1$.
- $\Omega_2 = \{ND\}$, dado que $X = 2$.
- $\Omega_3 = \{NND\}$, dado que $X = 3$, y así sucesivamente.

Por lo tanto, el rango de esta función es el conjunto $R_X = \{1, 2, 3, 4, \ldots\}$ o bien, $R_X = \mathbb{N}$.

En conclusión: Podemos ver que estos dos ejemplos tienen algo en común: sus rangos están formados por números enteros, lo cual hace que sean un tipo particular de variables aleatorias que reciben el nombre de **discretas**, las cuáles definiremos a continuación:

Definición: Una variable aleatoria se llama **variable aleatoria discreta** si se puede contar su conjunto de resultados posibles. Esto es equivalente a decir que toma valores en un conjunto "contable" o "numerable" de elementos.

Por ejemplo, el conjunto $\{-1, 1, 2, 3, 115, 1259\}$ es un conjunto numerable, al igual que lo es $\mathbb{N}$ que, aunque se trata de un conjunto infinito, también tiene la cualidad de que sus elementos pueden ser "enumerados". En los ejemplos anteriores, las variables aleatorias son discretas.

Distribución de probabilidad

Al igual que lo hemos hecho antes, arrancaremos con un ejemplo para construir ideas que nos permitirán comprender, de mejor manera, el concepto que definiremos aquí.

Ejemplo 3
Supongamos que lanzamos una moneda tres veces y anotamos en cada tirada si salió "cara" o "cruz".

ETIQUETA COMENTADO

El espacio muestral equiprobable de este experimento aleatorio es:

$$\Omega = \{CCC, CCX, CXC, CXX, XCC, XCX, XXC, XXX\}$$

Ahora, si definimos la variable aleatoria X como el número de caras,

¿con qué probabilidad toma cada uno de sus valores?

Según lo visto en el "Libro 1", la probabilidad en un espacio equiprobable la calculamos haciendo el cociente entre la "cantidad de casos favorables" y "el número de casos posibles". Resumimos esta información en la siguiente tabla:

x	0	1	2	3
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Por ejemplo, leemos la tabla anterior de la siguiente manera: la probabilidad de que X tome el valor 0, es decir, que no salga ninguna cara en los tres lanzamientos, es $P(X=0)=\frac{1}{8}$.

¿Qué sucede si sumamos todas las probabilidades?

$$P(X=0)+P(X=1)+P(X=2)+P(X=3)=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=1$$

Observen que los valores de x agotan todos los casos posibles. Por lo tanto, las probabilidades suman 1.

Una variable aleatoria discreta toma cada uno de sus valores con cierta probabilidad, como vimos en la actividad inicial.

Con frecuencia es conveniente representar todas las probabilidades de una variable aleatoria X usando una fórmula, la cual necesariamente sería una función de los valores numéricos x que denotaremos con $f(x), g(x), h(x)$ y así sucesivamente. Por lo tanto, escribimos $f(x)=P(X=x)$; es decir, $f(3)=P(X=3)$.

Al conjunto de pares ordenados $(x, f(x))$ se le llama **función de probabilidad** o **distribución de probabilidad** de la variable aleatoria discreta X.

Definición:

El conjunto de pares ordenados $(x, f(x))$ es una función de probabilidad o una distribución de probabilidad de la variable aleatoria discreta X si, para cada resultado x posible, se cumple:

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X=x)=f(x)$.

Ejemplo:

Acaba de llegar una importación de 30 celulares con iguales características de los cuales 5 vinieron con fallas técnicas de fábrica. Si una empresa compra al azar 3 de estos celulares, ¿Cuál es la distribución de probabilidad para el número de celulares con fallas técnicas?

Para resolver este ejemplo, vamos a definir a X como una variable aleatoria cuyos valores x son el número posible de celulares con fallas de fábrica comprados por la empresa. Entonces x sólo puede asumir los números $0, 1, 2$ y 3. Así, calculamos las siguientes probabilidades:

- $f(0) = P(X=0) = \frac{\binom{5}{0}\binom{25}{3}}{\binom{30}{3}} = \frac{115}{203} \approx 0,57$

- $f(1) = P(X=1) = \frac{\binom{5}{1}\binom{25}{2}}{\binom{30}{3}} = \frac{75}{203} \approx 0,37$

- $f(2) = P(X=2) = \frac{\binom{5}{2}\binom{25}{1}}{\binom{30}{3}} = \frac{25}{406} \approx 0,06$

- $f(3) = P(X=3) = \frac{\binom{5}{3}\binom{25}{0}}{\binom{30}{3}} = \frac{1}{406} \approx 0,002$

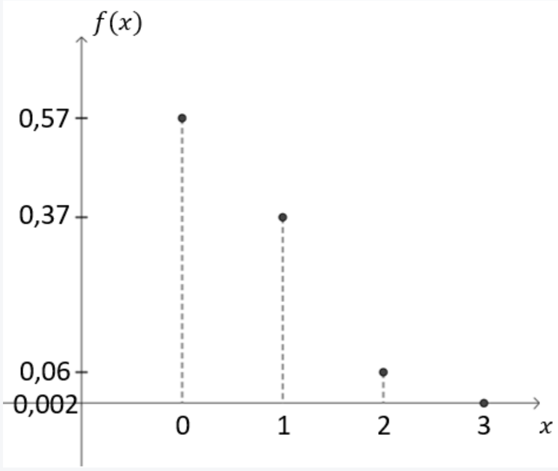
Por consiguiente, la distribución de probabilidad X es:

$$\begin{array}{ccccc} x & 0 & 1 & 2 & 3 \\ f(x) & \frac{115}{203} & \frac{75}{203} & \frac{25}{406} & \frac{1}{406} \end{array}$$

Podemos ver que la empresa tiene la mayor probabilidad de que ninguno de los tres celulares que compre vengan con fallas técnicas, ¿tiene sentido que sea así?

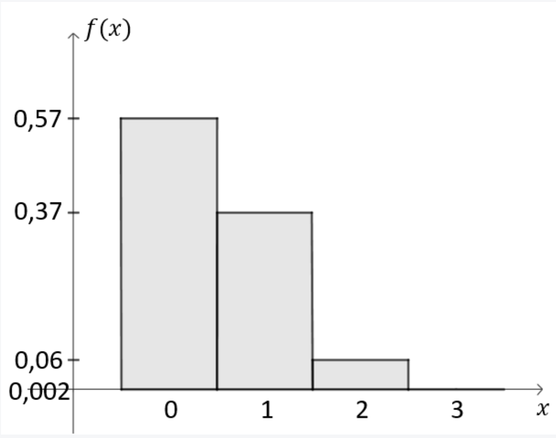
Representación gráfica:

A menudo, es útil ver una distribución de probabilidad en forma gráfica. Si en el eje horizontal colocamos los valores de la variable aleatoria X y en el vertical, la probabilidad de cada x, podemos graficar los puntos $(x, f(x))$ de nuestro ejemplo y obtendremos la siguiente gráfica:



Si unimos los puntos al eje x, ya sea con una línea punteada o con una línea sólida, obtenemos una gráfica de función de distribución de probabilidad. La figura anterior permite ver fácilmente qué valores de X tienen más probabilidad de ocurrencia.

En ocasiones, en vez de graficar los puntos $(x, f(x))$, lo que hacemos es construir rectángulos como se muestra a continuación:



Aquí los rectángulos se construyen de manera que sus bases, con la misma anchura, se centren en cada valor x, y que sus alturas igualen a las probabilidades correspondientes dadas por $f(x)$. Las bases se construyen de forma tal que no dejen espacios entre los rectángulos. Este gráfico se denomina **histograma de probabilidad.**

Observen que, como cada base de los rectángulos tienen el ancho de una unidad, $f(x)$ es igual al área del rectángulo centrado en x. Este concepto de utilizar áreas para representar probabilidades, es necesario para nuestro estudio de la distribución de probabilidad de una variable aleatoria continua.

Actividad inicial

¿Qué hubiera sucedido si en el ejemplo anterior nos piden calcular la probabilidad de que, en la compra de los celulares, como máximo vengan dos con fallas de fábrica?

Como estudiantes que ya han trabajado con Teoría de Conjuntos, saben que ahora debemos considerar la unión entre la posibilidad de que ninguno venga con fallas, uno solo o dos, es decir, $X \leq 2$, siendo X la variable aleatoria discreta que representa el número de celulares comprados con fallas de fábrica.

Por lo tanto, debemos calcular la probabilidad cuando $X \leq 2$, lo cual, podemos pensarlo como la suma de todas las probabilidades que cumplen esa condición, que fueron calculadas en la sección anterior, es decir:
$P(X \leq 2) = f(0) + f(1) + f(2) = \frac{115}{203} + \frac{75}{203} + \frac{25}{406} = \frac{405}{406}$

¿Tiene lógica el resultado no? Porque el único caso que estamos dejando fuera de análisis es $f(3) = \frac{1}{406}$, que es lo que le falta a lo anterior para llegar al total de 1.

Esto conduce a definir la función de la distribución acumulada de la variable aleatoria X.

Definición

La función de la distribución acumulada $F(x)$ de una variable aleatoria discreta X, con distribución de probabilidad $f(x)$ es
$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$ **para** $-\infty < x < \infty$

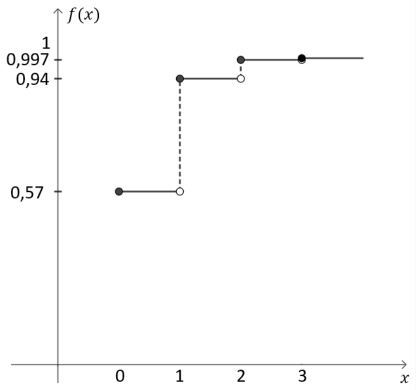
Para nuestro ejemplo, calculamos $F(2) = \frac{405}{406}$. De manera análoga, podemos encontrar la distribución acumulada de los demás valores de la variable aleatoria.

- $F(0) = f(0) = \frac{115}{203}$
- $F(1) = f(0) + f(1) = \frac{190}{203}$
- $F(2) = f(0) + f(1) + f(2) = \frac{405}{406}$
- $F(3) = f(0) + f(1) + f(2) + f(3) = 1$

Noten que para la distribución acumulada utilizamos la letra en mayúscula.

Representación gráfica:

Si graficamos la función de distribución acumulada del ejemplo, nos queda de la siguiente manera:



f(x)
1
0,997
0,94
0,57
0
1
2
3
x

¿Qué capacidades y habilidades vas a desarrollar?

El valor esperado y la varianza, son dos parámetros que -cuando existen- sirven para describir dónde y cómo se centra la distribución de probabilidad.

En esta sección, las definiremos, expondremos algunas propiedades útiles para operar con ellas y resolveremos algunos ejemplos.

¿Qué valor esperaríamos obtener?

Supongamos que tenemos dos monedas iguales, en las que de un lado tienen dibujado un cuadrado y en el otro, un triángulo.

Una persona lanzó estas dos monedas una cantidad de 16 veces y fue anotando los diferentes resultados, los cuales fueron:
$\{CC, TT, CC, TT, CT, TC, CC, CC, CT, TC, TC, TT, CT, CT, TT, TT\}$

Si definimos la variable aleatoria discreta X como el número de triángulos que resultan en cada lanzamiento, entonces los valores de X pueden ser 0, 1 y 2. A partir de esto y según el listado anterior, podemos armar la siguiente tabla:

Elementos del espacio muestral	Cantidad de defectuosos
0	4
1	7
2	5

Podríamos preguntarnos entonces lo siguiente: ¿cuál es el número promedio de triángulos por lanzamiento de las dos monedas?

Ya vimos en el curso de preingreso, que el promedio es la suma de todas las observaciones dividido por la cantidad de observaciones. Por lo tanto, planteamos y resolvemos:

$$\frac{0 \cdot 4 + 1 \cdot 7 + 2 \cdot 5}{16} = 1,0625$$

Éste es un valor promedio de los datos, aunque no es un resultado posible de $\{0, 1, 2\}$. Por lo tanto, un promedio no es necesariamente un resultado posible del experimento.

Reestructuremos ahora nuestro cálculo del número promedio de caras para tener las siguientes expresiones equivalentes:

$$\frac{0 \cdot 4 + 1 \cdot 7 + 2 \cdot 5}{16} = 1,0625$$

Los números $\frac{4}{16}$, $\frac{7}{16}$ y $\frac{5}{16}$ corresponden a la distribución de probabilidad de los diferentes valores de X en nuestro experimento, es decir, $f(0)$, $f(1)$ y $f(2)$. Tales fracciones también son llamadas **frecuencias relativas**.

Por lo tanto, el cálculo anterior lo podemos reescribir como:
$$0 \cdot f(0) + 1 \cdot f(1) + 2 \cdot f(2) = 1,0625$$

Lo anterior no es más que la suma de los productos entre el valor de la variable aleatoria discreta X y su probabilidad correspondiente.

Entonces, realmente podemos calcular la media o el promedio de un conjunto de datos, si conocemos los distintos valores que ocurren y sus frecuencias relativas, sin tener conocimiento del número total de observaciones en el conjunto de datos.

A este valor promedio se le conoce como **media de la variable aleatoria** X o **media de la distribución de probabilidad** de X, y se le denota como μ_x o simplemente como μ cuando es evidente a qué variable aleatoria se está haciendo referencia. También es común referirse a esta media como la **esperanza matemática** o el **valor esperado** de la variable aleatoria X y denotarla como $E(X)$.

¿Por qué se llamará de esta manera? Siguiendo con nuestro ejemplo, observen que el valor esperado dio $1,0625$, es decir, está muy cercano al valor 1. Esto quiere decir que si realizamos un próximo lanzamiento, uno esperaría obtener un solo triángulo en las caras de las dos monedas.

Definición

Sea X una variable aleatoria discreta con distribución de probabilidad $f(x)$. La media o valor esperado de X es

$$\mu = E(X) = \sum_{x} x \cdot f(x)$$

Ejemplo

En un grupo de 7 estudiantes, hay 4 que van a un instituto de inglés y 3 que no van. Un profesor selecciona al azar una muestra de 3 estudiantes. Calcular el valor esperado del número de estudiantes que van a un instituto de inglés.

Sea X el número de estudiantes que van a un instituto de inglés. La función de distribución de probabilidad de X es:

$$f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}, \text{ con } x = 0, 1, 2, 3$$

Haciendo cálculos como ya se hizo antes, tenemos que $f(0) = \frac{1}{35}, f(1) = \frac{12}{35}, f(2) = \frac{18}{35}$ y $f(3) = \frac{4}{35}$. Por lo tanto,

$$\mu = E(X) = 0 \cdot \frac{1}{35} + 1 \cdot \frac{12}{35} + 2 \cdot \frac{18}{35} + 3 \cdot \frac{4}{35} = \frac{12}{7} \approx 1,714.$$

De esta manera, en promedio esperaríamos que en la muestra haya aproximadamente dos estudiantes que van a un instituto de inglés.

La esperanza o valor esperado, tiene algunas propiedades o teoremas que son importantes conocerlos. Aquí solamente enunciaremos el siguiente y resolveremos un ejemplo para comprender de qué trata:

Enunciación del teorema:

Sea X una variable aleatoria discreta con distribución de probabilidad $f(x)$. El valor esperado de la variable aleatoria $g(X)$ es

$$\mu_{g(X)} = E[g(X)] = \sum_{x} g(x)f(x)$$

La variable aleatoria $g(X)$ depende de X, es decir, cada valor de $g(X)$ es determinado por el valor de X.

Ejemplo:

Suponga que el número de automóviles X que pasa por un local de lavado de autos entre las 4:00 p.m. y las 5:00 p.m. de cualquier viernes soleado, tiene la siguiente distribución de probabilidad:

x	4	5	6	7	8	9
$f(x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Sea $g(X) = 1000X$ la cantidad de dinero que el administrador paga al operador. Calcular las ganancias esperadas del operador en este periodo específico.

Vemos que tenemos una variable aleatoria $g(X)$ que depende de la variable aleatoria X. Según el teorema anterior, podemos plantear lo siguiente:

$$E[g(X)] = E(1000X) = \sum_{x=4}^{9} (1000x)f(x)$$

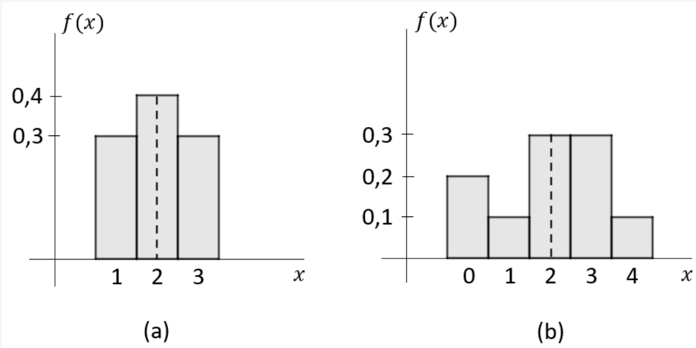
$$E[g(X)] = (1000 \cdot 4) \cdot \frac{1}{12} + (1000 \cdot 5) \cdot \frac{1}{12} + (1000 \cdot 6) \cdot \frac{1}{4} + (1000 \cdot 7) \cdot \frac{1}{4} + (1000 \cdot 8) \cdot \frac{1}{6} + (1000 \cdot 9) \cdot \frac{1}{6}$$

$$E[g(X)] = 6833,33$$

Por lo tanto, el operador esperaría obtener una ganancia de $\$6833,33$.

¿Cuál distribución tiene más variabilidad?

Ya vimos que la media o valor esperado de una variable aleatoria X es de especial importancia, porque describe en dónde se centra la distribución de probabilidad. Sin embargo, esta por si sola no ofrece una descripción adecuada de la forma de la distribución. También se necesita clasificar la variabilidad en la distribución. En la imagen de abajo tenemos dos distribuciones de probabilidad discretas con la misma media ($\mu = 2$), pero que difieren de manera considerable en la variabilidad o dispersión de sus observaciones sobre la media.



La medida de variabilidad más importante de una variable aleatoria X se denomina **varianza** y en general se denota con el símbolo σ^2.

Definición

Sea X una variable aleatoria discreta con distribución de probabilidad $f(x)$ y media μ. La varianza de X es

$$\sigma^2=E[(X- \mu)^2]= \displaystyle\sum_{x}(x- \mu)^2 f(x)$$

La raíz cuadrada positiva de la varianza, σ, se llama **desviación estándar** de X.

La cantidad $x- \mu$ en la definición se llama **desviación de una observación** respecto a su media. Como estas desviaciones se elevan al cuadrado y después se promedian, σ^2 será mucho menor para un conjunto de valores x que estén cercanos a μ, que para un conjunto de valores que varíe de forma considerable de μ.

Ejemplo (utilizando los gráficos del principio)

Suponga que la variable aleatoria X representa el número de pedidos cancelados en un día. La distribución de probabilidad para la empresa A es:

x	1	2	3
$f(x)$	$0,3$	$0,4$	$0,3$

y para el empresa B es:

x	0	1	2	3	4
$f(x)$	$0,2$	$0,1$	$0,3$	$0,3$	$0,1$

Vamos a probar que la varianza de la distribución de probabilidad para la empresa B es mayor que la de la empresa A.

Primero, para la empresa A encontramos que:

$$\mu_A=E(X)=1 \cdot 0,3 + 2 \cdot 0,4 + 3 \cdot0,3 = 2$$

y entonces

$$\sigma_A ^2 = \displaystyle\sum_{x=1}^{3} (x-2)^2 f(x)=(1-2)^2 \cdot 0,3 + (2-2)^2 \cdot 0,4+(3-2)^2 \cdot 0,3 = 0,6$$

Por su parte, para el empresa B tenemos:

$$\mu_B = E(X)=0 \cdot0,2+1 \cdot0,1+2 \cdot0,3+3 \cdot0,3+4 \cdot0,1=2$$

y entonces

$$\sigma_B ^2 = \displaystyle\sum_{x=0}^{4} (x-2)^2 f(x)=(0-2)^2 \cdot 0,2 + (1-2)^2 \cdot 0,1+(2-2)^2 \cdot 0,3 + (3-2)^2 \cdot 0,3 + (4-2)^2 \cdot 0,1 = 1,6$$

Observen que hemos comprobado lo que ya habíamos anticipado mirando solamente los histogramas.

Al igual que la esperanza, la varianza tiene teoremas importantes. A continuación, enunciaremos dos de ellos:

Enunciación del teorema 1:

Una fórmula alternativa que se prefiere para calcular $\sigma ^2$, que a menudo simplifica los cálculos, se establece en el siguiente teorema.

La varianza de una variable aleatoria X es:

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

Se deja como actividad que calculen las varianzas anteriores empleando la fórmula de este teorema y validen que funciona.

Enunciación del teorema 2:

Sea X una variable aleatoria discreta con distribución de probabilidad $f(x)$. La varianza de la variable aleatoria $g(X)$ es

$$\sigma^2_{g(X)}=E \lbrace{[g(X)- \mu_{g(X)}]^2}\rbrace = \displaystyle\sum_{x} [g(x)- \mu_{g(X)}]^2 f(x)$$

Teorema 3

La covarianza de dos variables aleatorias X e Y, con medias μ_x y μ_y, respectivamente, está dada por

$$\sigma_{XY}=E(XY)- \mu_x \mu_y$$

Definición

Sean X e Y variables aleatorias con covarianza σ_{XY} y desviaciones estándar σ_{X} y σ_{Y}, respectivamente. El coeficiente de correlación de X y Y es

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

¿Qué capacidades y habilidades vas a desarrollar?

A menudo, las observaciones que se generan mediante diferentes experimentos aleatorios tienen el mismo tipo general de comportamiento. En consecuencia, las variables aleatorias discretas asociadas con estos experimentos se pueden describir esencialmente con la misma distribución de probabilidad y, por lo tanto, es posible representarlas usando una sola fórmula.

En esta sección se estudiarán distribuciones discretas especiales que se aplican a diversas situaciones experimentales. De hecho, se necesitan solo unas cuantas distribuciones de probabilidad importantes para describir muchas de las variables aleatorias discretas que se encuentran en la práctica.

¿Qué capacidades y habilidades vas a desarrollar?

Con frecuencia un experimento consta de pruebas repetidas, cada una con dos resultados posibles que se pueden denominar éxito o fracaso. Por ejemplo:

Éxito	Fracaso
Que salga cara	Que salga cruz
Que acierte un lanzamiento	Que falle un lanzamiento
Que salga el 1	Que salga cualquier otro número
Que un producto no esté defectuoso	Que sea defectuoso

Podemos elegir definir cualquiera de los resultados como éxito. El proceso se conoce como proceso de Bernoulli y cada ensayo se denomina experimento de Bernoulli.

El proceso de Bernoulli

En términos estrictos, el proceso de Bernoulli se caracteriza por lo siguiente:

1. El experimento consta de n ensayos repetidos.
2. Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso.
3. La probabilidad de un éxito, que se denota con p, permanece constante de un ensayo a otro.
4. Los ensayos repetidos son independientes.

Ejemplo

Un jugador de baloncesto realizó 20 tiros libres, de los cuales acertó 16 y falló 4. Si tira tres lanzamientos seguidos, ¿cuál es la probabilidad de que acierte dos de ellos?

En la siguiente página resolveremos este problema, lo que haremos aquí es analizar si se adecúa o no con un proceso de Bernoulli. Analicemos cada una de las condiciones:

1. En este caso un ensayo sería lanzar la pelota y observar su resultado. Como se repite tres veces y los lanzamientos serán con las mismas condiciones, cumple este primer pedido.
2. Cada lanzamiento va a producir como resultado un éxito (que sería acertar) o un fracaso (que sería no acertar). Por lo que también cumple la segunda condición.
3. La probabilidad de éxito permanece constante de un ensayo a otro, la cual se puede calcular fácilmente como veremos luego.
4. Los ensayos repetidos son independientes, dado que el resultado de un lanzamiento no condiciona el del otro.

Por lo tanto, como este experimento cumple las cuatro condiciones es un **proceso de Bernoulli**.

¿Qué capacidades y habilidades vas a desarrollar?

Actividad inicial:

Ahora si vamos a resolver el ejemplo de la sección anterior. Volvamos a leer su enunciado:
Un jugador de baloncesto realizó 20 tiros libres, de los cuales acertó 16 y falló 4. Si tira tres lanzamientos seguidos, ¿cuál es la probabilidad de que acierte dos de ellos?

Llamaremos p a la probabilidad de éxito, es decir, a que acierte un tiro libre y q a la probabilidad de que no acierte. Entonces, tenemos que $p= \frac{16}{20}=0,8$ y $q=1-p=0,2$

Si indicamos con E al tiro acertado y con F al errado, puede pasar una de las siguientes posibilidades:

1° lanzamiento	2° lanzamiento	3° lanzamiento
E	E	F
E	F	E
F	E	E

Como los sucesos son independientes, dado que si erra el primero nada implica lo que vaya a pasar con el segundo, podemos calcular las siguientes probabilidades aplicando la propiedad de la independencia de eventos: QUE HACER CON TODAS LAS FÓRMULAS DE ABAJO

$P(E \cap E \cap F)=P(E) \cdot P(E) \cdot P(F)=0,8 \cdot 0,8 \cdot 0,2=0,128$

$P(E \cap F \cap E)=P(E) \cdot P(F) \cdot P(E)=0,8 \cdot 0,2 \cdot 0,8=0,128$

$P(F \cap E \cap F)=P(F) \cdot P(E) \cdot P(E)=0,2 \cdot 0,8 \cdot 0,8=0,128$

Sumamos las tres cantidades anteriores, o bien, hacemos: $3 \cdot 0,128=0,384$.

Por lo tanto, la probabilidad de que acierte dos de los tres tiros es de $0,384$.

Observarán que el cálculo de todas las probabilidades anteriores es el mismo, lo único que cambia es el orden, pero como la multiplicación es conmutativa, el orden de los factores no altera el producto.

Ahora bien, qué hubiera sucedido si nos preguntan: ¿cuál es la probabilidad de que acierte dos tiros libres si lanza cuatro veces?

Hacemos una tabla similar a la anterior y nos queda:

1° lanzamiento	2° lanzamiento	3° lanzamiento	4° lanzamiento
E	E	F	F
E	F	E	F
E	F	F	E
F	E	E	F
F	E	F	E

F	F	E	E

Vimos que todas las probabilidades se calculan de la misma manera, entonces calculamos una sola y lo multiplicamos por 6 que es la cantidad de posibilidades:

$6 \cdot P(E) \cdot P(E) \cdot P(F) \cdot P(F)=6 \cdot 0,8 \cdot 0,8 \cdot 0,2 \cdot 0,2 = 0,1536$

A lo anterior, también podríamos indicarlo como: $6 \cdot 0,8^2 \cdot 0,2^2$

Remarquemos algunas cuestiones:

- El valor de la probabilidad de éxito y fracaso la conocemos.
- El exponente del valor de cada probabilidad, coincide con el número de éxitos o fracasos, respectivamente.

Toda esta información la tenemos, lo difícil es determinar cuántas posibilidades vamos a tener. Pero esto lo resolvemos utilizando combinatoria con repetición. En el último caso los elementos son p,p,q,q, es decir, son 4 elementos tomados de a 2 que se resuelve mediante el número combinatorio:

$$\binom{4}{2} = \frac{4!}{2! \cdot 2!} = 6$$

lo cual obviamente coincide con lo que habíamos calculado antes.

Entonces, con esta generalización es más simple resolver este tipo de problemas. Por ejemplo, si nos preguntaran: ¿cuál es la probabilidad de que acierte tres tiros de cinco lanzamientos?

Calculamos el número combinatorio y lo multiplicamos por la probabilidad de tener tres aciertos y dos fallos:

$$\binom{5}{3} 0,8^3 \cdot 0,2^2 = 0,2048$$

Este ejemplo corresponde a una **distribución binomial**.

El número X de éxitos en n experimentos de Bernoulli se denomina **variable aleatoria binomial**. La distribución de probabilidad de esta variable aleatoria discreta se llama **distribución binomial.**

Distribución binomial

Un experimento de Bernoulli puede tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 – p$. Entonces, la distribución de probabilidad de la variable aleatoria binomial X, el número de éxitos en n ensayos independientes, es

$$b(x,n,p)= \binom{n}{x} p^xq^{n-x} \text{ con } x=0,1,2,...,n$$

Dicho esto, retomemos nuestra actividad inicial y vamos a analizar lo que sucede con 7 lanzamientos.

Ya dijimos que corresponde a una distribución binomial, porque cumple con las condiciones de un proceso de Bernoulli. Entonces:

$$n=7, \quad p=0,8$$

y la fórmula de la función de distribución binomial es:

$$b(x;7;0,8)= \binom{7}{x}0,8^x \cdot 0,2^{7-x}$$

Por lo tanto, si queremos analizar la probabilidad de que acierte 5 de esos 7 lanzamientos, tomamos la fórmula anterior y hacemos que $x=5$, como se muestra a continuación:

$$b(5;7;0,8)= \binom{7}{5}0,8^5 \cdot 0,2^{7-5} \approx 0,28$$

Media y varianza

La media y la varianza de la distribución binomial $b(x,n,p)$ son:

$$\mu=np \quad \text{y} \quad \sigma^2=npq$$

En nuestro ejemplo, sería:

$$\mu=7 \cdot 0,8=5,6 \quad \text{y} \quad \sigma^2=7 \cdot 0,8 \cdot 0,2=1,12$$

Experimentos binomiales negativos

Consideremos un experimento con las mismas propiedades de un experimento binomial, solo que en este caso las pruebas se repetirán hasta que ocurra un número fijo de éxitos. Por lo tanto, en vez de encontrar la probabilidad de x éxitos en n pruebas, donde n es fija, ahora nos interesa la probabilidad de que ocurra el $k-$ésimo éxito en la $x-$ésima prueba. Los experimentos de este tipo se llaman **experimentos binomiales negativos**.

Volvamos a la idea de la detección de los correos no deseados, ¿la recuerdas?
Supongamos que el algoritmo que diseñamos tiene un 60% de eficacia. El uso del algoritmo será exitoso cuando logre detectar alguna anomalía sobre el correo. Nos interesa calcular la probabilidad de que el quinto correo no deseado sea el séptimo. Si designamos un éxito con E y un fracaso con F, un orden posible para alcanzar el resultado que se desea es $EFEEEFE$, que ocurre con la siguiente probabilidad:

$$0,6 \cdot 0,4 \cdot 0,6 \cdot 0,6 \cdot 0,6 \cdot 0,4 \cdot 0,6=0,6^5 \cdot0,4^2$$

Podríamos listar todos los posibles ordenamientos reacomodando las E y las F, con excepción del último resultado, que debe ser el quinto éxito. El número total de ordenamientos posibles se puede calcular haciendo:

$$\binom {6} {4} = 15$$

Por lo tanto, si X representa el resultado en el que ocurre el quinto éxito, entonces

$$P(X=7)= \binom {6} {4}0,6^50,4^2=0,1866$$

Ahora bien, analicemos la siguiente pregunta: ¿cuál es la probabilidad de que el tercer correo no deseado sea el décimo?

Un orden posible para alcanzar el resultado que se desea es $FEFFFEFFFE$, que ocurre con la siguiente probabilidad:

$$0,4 \cdot 0,6 \cdot 0,4 \cdot 0,4 \cdot 0,4 \cdot 0,6 \cdot 0,4 \cdot 0,4 \cdot 0,4 \cdot 0,6= 0,4^7 \cdot 0,6^3 \approx 0,00035$$

Para encontrar todos los posibles ordenamientos reacomodando las E y las F, hacemos:

$$\binom {9} {2}=36$$

Por lo tanto, si X representa el resultado en el que ocurre el tercer éxito, entonces

$$P(X=10)= \binom {9} {2}0,6^3 \cdot 0,4^7=0,0127$$

Si observamos con detenimiento esta última fórmula (junto con la anterior) la podemos reescribir así:

$$P(X=10)= \binom {10-1} {3-1}0,6^3 \cdot 0,4^{10-3}$$

Lo cual da origen a la **función de distribución binomial negativa**.

Distribución binomial negativa

Si ensayos independientes repetidos pueden dar como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 – p$, entonces la distribución de probabilidad de la variable aleatoria X, el número del ensayo en el que ocurre el $k-$ésimo éxito, es

$$b^{*}(x,k,p)= \binom {x-1} {k-1}p^kq^{x-k} \text{ con } x=k,k+1,k+2,...$$

Distribución geométrica

Si consideramos el caso especial de la distribución binomial negativa, donde $k=1$, tenemos una distribución de probabilidad para el número de ensayos que se requieren para un solo éxito. Un ejemplo sería lanzar un dado hasta que salga un seis. Nos podemos interesar en la probabilidad de que el primer seis resulte en el quinto lanzamiento. En este caso la distribución binomial negativa se reduce a la forma

$$b^{*}(x,1,p)=pq^{x-1} \text{ con } x=1,2,3,...$$

Como los términos sucesivos constituyen una progresión geométrica, se acostumbra referirse a este caso especial como **distribución geométrica** y denotar sus valores con $g(x, p)$.

Si pruebas independientes repetidas pueden tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X, el número de la prueba en el que ocurre el primer éxito, es

$$g(x,p)=pq^{x-1} \text{ con } x=1,2,3,...$$

Media y varianza

La media y la varianza de una variable aleatoria que sigue la distribución geométrica son:

$$\mu = \frac{1}{p} \quad \text{y} \quad \sigma ^2= \frac{1-p}{p^2}$$

Ejemplo

Supongamos que nos interesa averiguar la probabilidad de que al lanzar una moneda, obtengamos la primera cara en el segundo lanzamiento.

En este caso estamos en una distribución geométrica, ¿te diste cuenta por qué? Porque nos interesa averiguar únicamente el primer éxito (que salga cara), el cual se pretende lograr en el segundo lanzamiento.

La probabilidad de éxito es $0,5$, al igual que la del fracaso. Por lo tanto, la fórmula de la función nos quedaría:

$$g(x;0,5)=0,5 \cdot 0,5^{x-1} \text{ con } x=1,2,3,...$$

Como nos interesa que sea en el segundo lanzamiento, $x=2$ y nos queda:

$$g(2;0,5)=0,5 \cdot 0,5^{2-1}=0,25$$

Si calculamos su media y varianza, tenemos:

$$\mu = \frac{1}{0,5}=2 \quad \text{y} \quad \sigma ^2= \frac{1-0,5}{0,5^2} = 2$$

No siempre va a suceder que ambas medidas coincidan.

Proceso de Poisson

Los experimentos que producen valores numéricos de una variable aleatoria X, en donde el número de resultados ocurren durante un intervalo de tiempo determinado o en una región específica, se denominan **experimentos de Poisson**. Por ejemplo, un experimento de Poisson podría generar observaciones para la variable aleatoria X que representa:

- el número de llamadas telefónicas por hora que recibe una oficina,
- el número de juegos suspendidos debido a algún incidente durante la temporada de fútbol,
- el número de bacterias en un cultivo dado,
- el número de píxeles en una determinada pantalla.

Propiedades del proceso de Poisson

El proceso de Poisson tiene las siguientes propiedades:

1. El número de resultados que ocurren en un intervalo o región específica, es independiente del número que ocurre en cualquier otro intervalo de tiempo o región del espacio disjunto. A partir de esto, se dice que el proceso de Poisson no tiene memoria.
2. La probabilidad de que ocurra un solo resultado durante un intervalo de tiempo muy corto o en una región pequeña es proporcional a la longitud del intervalo o al tamaño de la región, y no depende del número de resultados que ocurren fuera de este intervalo de tiempo o región.
3. La probabilidad de que ocurra más de un resultado en tal intervalo de tiempo corto o que caiga en tal región pequeña es insignificante.

El número X de resultados que ocurren durante un experimento de Poisson se llama **variable aleatoria de Poisson** y su distribución de probabilidad se llama **distribución de Poisson**.

Distribución de Poisson

La distribución de probabilidad de la variable aleatoria de Poisson X, la cual representa el número de resultados que ocurren en un intervalo de tiempo dado o región específicos y se denota con t, es

$$p(x, \lambda t)= \frac{e^{- \lambda t} (\lambda t)^x}{x!} \text{ con } x=0,1,2,3,...$$

donde λ es el número promedio de resultados por unidad de tiempo, distancia, área o volumen y e es el número irracional cuyas primeras cifras decimales son $e=2,71828...$

Excede a este curso la deducción de la fórmula para la distribución de Poisson.

Ejemplo 1

En una pantalla de $1m^2$, en promedio fallan 5 LED. ¿Cuál es la probabilidad de que fallen 7 LED en una pantalla de $1m^2$?

Si analizamos con detenimiento, podremos ver que las tres propiedades descritas anteriormente se cumplen para ser considerado un proceso de Poisson.

Luego, identificamos lo necesario: $x=7$ y $\lambda t = 5$. Reemplazamos en la fórmula de la distribución de Poisson y calculamos:

$$p(7,5)= \frac{e^{- 5} 5^7}{7!} \approx 0,104$$

Por lo tanto, la probabilidad de que fallen 7 LED en una pantalla de ese tamaño es de aproximadamente $0,104$.

Ejemplo 2:

Una persona desarrolló un app que calcula el promedio de mensajes que se reciben en un determinado momento y la instaló en el celular de una amiga que tiene un emprendimiento.

Con ella pudo determinar que, por hora, en promedio le escriben cuatro personas. ¿Cuál es la probabilidad de que en la siguiente hora le escriban más de tres personas?

Sea X el número de mensajes que recibe sobre su emprendimiento. Entonces, debemos calcular la siguiente probabilidad: $P(X>3)=1-P(X \leq 3)$.

Como se trata de una distribución de Poisson, hacemos lo siguiente:

$P(x>3)=1-P(x \leq 3) = 1-\displaystyle\sum_{x=0}^{3} p(x,4) \approx 0,567$

Por lo tanto, la probabilidad de que en la siguiente hora le escriban más de tres personas es de aproximadamente $0,567$.

Media y varianza

Tanto la media como la varianza de la distribución de Poisson $p(x, \lambda t)$ son λt, donde t es el tiempo, la distancia, el área o el volumen específicos de interés.

En el ejemplo 2, si queremos calcular la media en 6 horas, hacemos: $\mu = 4 \cdot 6 =24$. Lo mismo con la varianza.

Actividad inicial:

En cierta parte del país, 1 de cada 50 niños se ha roto accidentalmente un hueso del cuerpo a la edad de 5 años. Si se eligen al azar 400 niños de 5 años, ¿cuál es la probabilidad de que exactamente 3 de ellos hayan sufrido una fractura?

Observemos que este problema lo podemos resolver con una distribución binomial, ¿se dieron cuenta por qué?

Definimos a nuestra variable aleatoria X como el número de niños que ha sufrido una fractura. Además, la probabilidad de éxito es $p= \frac{1}{50}$ y la de fracaso es $q= \frac{49}{50}$. Luego, aplicamos la fórmula y calculamos su probabilidad:

$$P(X=3)= \binom {400} {3} \cdot (\frac{1}{50})^3 \cdot (\frac{49}{50})^{397}$$

Sin embargo, cuando queremos calcularlo con la calculadora nos marca error, eso se debe a que las operaciones dan números muy grandes (o muy chicos) y no está diseñada para operar con tales números.

Si usamos algún programa más potente o trabajamos de forma algebraica antes, vemos que el resultado es: $P(X=3) \approx 0{,}02783599473$

Ahora conviene preguntarnos: ¿no podremos resolver este problema empleando otra distribución? La respuesta es sí y se trata de la distribución de Poisson.

Si se fijan con atención, debería ser evidente que estas dos distribuciones están relacionadas.

En el caso de la distribución binomial, si n es bastante grande y p es pequeña (como nuestra actividad inicial), las condiciones comienzan a simular las implicaciones de espacio o tiempo continuos del proceso de Poisson. Esto motiva a enunciar el siguiente teorema:

Teorema

Sea X una variable aleatoria binomial con distribución de probabilidad $b(x, n, p)$. Cuando $n \to \infty, p \to 0$, y $np\to \mu$ permanece constante, entonces:

$$b(x,n,p)\to p(x, \mu) \text{ cuando } n \to \infty$$

Volviendo a nuestra actividad inicial...

Si ahora utilizamos una distribución de Poisson para resolver la probabilidad, nos queda:

$P(X=3)= \frac{e^{-8} \cdot 8^3 }{3!} \approx 0{,}02862614425$

donde λ lo calculamos haciendo $\lambda=n \cdot p=400 \cdot \frac{1}{50}=8$.

Si comparan la probabilidad encontrada con la distribución binomial y la de Poisson, ambas están muy cercanas, por eso es una aproximación.