Curso introductorio en Matematicas y Estadistica

Programa de Finanzas Cuantitativas

QUANt

UNIVERSIDAD DEL CEMA

MCMLXXVIII

UCEMA

Probabilidad

Concepto de probabilidad - Probabilidad discreta

En la teoría de probabilidad la aleatoriedad se modela de la siguiente forma:

- Con un **espacio muestral** S que contiene todos los posibles estados que puede tomar el universo en el que queremos hacer mediciones.

Concepto de probabilidad - Probabilidad discreta

En la teoría de probabilidad la aleatoriedad se modela de la siguiente forma:

- Con un **espacio muestral** S que contiene todos los posibles estados que puede tomar el universo en el que queremos hacer mediciones.

- Con una colección F de subconjuntos de S a los que **se les puede asignar una probabilidad** de ocurrir.

Concepto de probabilidad - Probabilidad discreta

En la teoría de probabilidad la aleatoriedad se modela de la siguiente forma:

- Con un **espacio muestral** *S* que contiene todos los posibles estados que puede tomar el universo en el que queremos hacer mediciones.

- Con una colección *F* de subconjuntos de *S* a los que **se les puede asignar una probabilidad** de ocurrir.

- Con una **función de probabilidad** *P* que asigna a cada subconjunto de *S* presente en *F* un valor entre 0 y 1 que representa su probabilidad de ocurrir.

Concepto de probabilidad - Probabilidad discreta

En la teoría de probabilidad la aleatoriedad se modela de la siguiente forma:

- Con un **espacio muestral** *S* que contiene todos los posibles estados que puede tomar el universo en el que queremos hacer mediciones.

- Con una colección *F* de subconjuntos de *S* a los que **se les puede asignar una probabilidad** de ocurrir.

- Con una **función de probabilidad** *P* que asigna a cada subconjunto de *S* presente en *F* un valor entre 0 y 1 que representa su probabilidad de ocurrir.

Importante: para definir un espacio de probabilidad, *S, F y P* deben cumplir ciertas propiedades. Fuera de esto, las elecciones son completamente arbitrarias, y cada elección representará una forma distinta de modelar el problema que nos interese.

Concepto de probabilidad

Las propiedades que deben cumplirse son:

F1. $S \in F$.

F2. Si $A \in F \Rightarrow A^c \in F$.

F3. Si $A_i \in F \Rightarrow \bigcup_{i=0}^{\infty} A_i \in F$.

P1. $P(S) = 1$.

P2. Si $A_i \in F$ son disjuntos entonces $P\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} P(A_i)$.

El trio (*S*, *F*, *P*) se llama **espacio de probabilidad**.

Concepto de probabilidad - Espacios discretos - Ejemplo

Supongamos que lanzamos dos dados y queremos hacer mediciones sobre los resultados.

Concepto de probabilidad - Espacios discretos - Ejemplo

Supongamos que lanzamos dos dados y queremos hacer mediciones sobre los resultados.

Podemos modelar como:

$$
\begin{aligned}
S = \{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\
&(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\
&(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\
&(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\
&(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\
&(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}
\end{aligned}
$$

y $P((i,j)) = \dfrac{1}{36}$.

Concepto de probabilidad - Espacios discretos - Ejemplo

Supongamos que lanzamos dos dados y queremos hacer mediciones sobre los resultados.

Podemos modelar como:

$$\begin{aligned} S = \{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\ &(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\ &(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\ &(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\} \end{aligned}$$

y $P((i,j)) = \dfrac{1}{36}$.

Un espacio de probabilidad con *S* finito* se llama **discreto**.

**O con un subconjunto finito de probabilidad 1.*

Concepto de probabilidad - Espacios discretos - Ejemplo

Se pueden responder preguntas sobre la probabilidad de que sucedan distintos **eventos**:

a. Cual es la probabilidad de sacar un 6 y un 2, en ese orden?
b. Cual es la probabilidad de sacar un 6 y un 2?
c. Cual es la probabilidad de sacar dos 5?
d. Cual es la probabilidad de sacar dos números iguales?
e. Cual es la probabilidad de que los números elegidos sumen 7?
f. Cual es la probabilidad de que los números elegidos sumen 2?
g. Cual es la probabilidad de sacar un número par y un número impar?

Notación: usualmente denotaremos a los eventos con una letra. Ejemplo: si A = {los dados suman 7} entonces P(A) = 6/36 = 1/6.

Concepto de probabilidad - Espacios continuos

Supongamos que queremos entender la temperatura en el mundo.

Podríamos definir $S = \mathbb{R}$ el espacio de posibles temperaturas.

Concepto de probabilidad - Espacios continuos

Supongamos que queremos entender la temperatura en el mundo.

Podríamos definir $S = \mathbb{R}$ el espacio de posibles temperaturas.

También podríamos modelar $S = \boxed{(\mathbb{R} \times \mathbb{R})} \times \boxed{\mathbb{R}_{>0}} \times \boxed{\mathbb{R}}$.

Coordenadas (espacio) Tiempo Temperatura

Concepto de probabilidad - Espacios continuos

Supongamos que queremos entender la temperatura en el mundo.

Podríamos definir $S = \mathbb{R}$ el espacio de posibles temperaturas.

También podríamos modelar $S = \underset{\text{Coordenadas (espacio)}}{\boxed{(\mathbb{R} \times \mathbb{R})}} \times \underset{\text{Tiempo}}{\boxed{\mathbb{R}_{>0}}} \times \underset{\text{Temperatura}}{\boxed{\mathbb{R}}}$.

En este caso, cuál debería ser la probabilidad de observar una temperatura particular?

Un espacio que no es discreto se llama **continuo**.

Probabilidad condicional - Definición

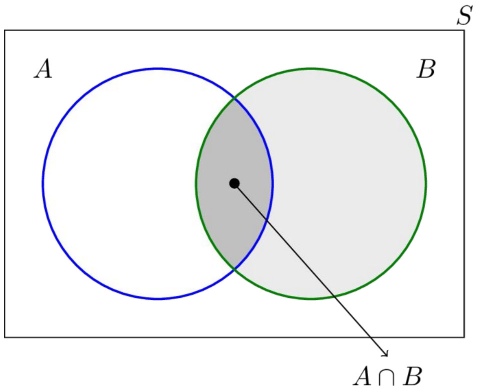
Un concepto importante es la **probabilidad condicional**. Es una forma de medir la probabilidad de que ocurra un evento dado que otro ocurrió.

Probabilidad condicional - Definición

Un concepto importante es la **probabilidad condicional**. Es una forma de medir la probabilidad de que ocurra un evento dado que otro ocurrió.

Formalmente se define como:

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



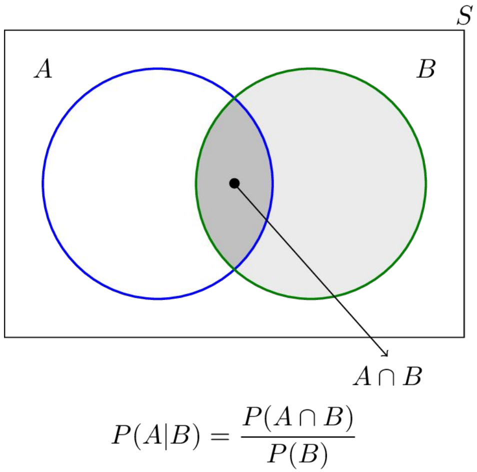
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Probabilidad condicional - Definición

Un concepto importante es la **probabilidad condicional**. Es una forma de medir la probabilidad de que ocurra un evento dado que otro ocurrió.

Formalmente se define como:

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



Intuitivamente es el resultado de restringir el espacio muestral a B y medir la probabilidad de que A ocurra allí dentro. Un resultado intuitivo es el siguiente:

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

Probabilidad condicional - Ejemplos

En el ejemplo de los dos dados:

- Cual es la probabilidad de que los dados sumen 6 dado que ambos son impares?

Probabilidad condicional - Ejemplos

En el ejemplo de los dos dados:

- Cual es la probabilidad de que los dados sumen 6 dado que ambos son impares?

    - A = {los dados suman 6} = {(1,5), (2,4), (3,3), (4,2), (5,1)}
    - B = {ambos dados son impares} = {...}
    - $A \cap B$ = {los dados suman 6 y son impares} = {(1,5), (3,3), (5,1)}

Probabilidad condicional - Ejemplos

En el ejemplo de los dos dados:

- Cual es la probabilidad de que los dados sumen 6 dado que ambos son impares?

  - A = {los dados suman 6} = {(1,5), (2,4), (3,3), (4,2), (5,1)}
  - B = {ambos dados son impares} = {...}
  - A∩B = {los dados suman 6 y son impares} = {(1,5), (3,3), (5,1)}

Entonces $P(A) = \frac{5}{36}$, $P(B) = \frac{9}{36}$, $P(A \cap B) = \frac{3}{36}$, luego:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/36}{9/36} = \frac{1}{3}$$

Probabilidad condicional - Ejemplos

En el ejemplo de los dos dados:

- Cual es la probabilidad de que los dados sumen 6 dado que ambos son impares?

    - A = {los dados suman 6} = {(1,5), (2,4), (3,3), (4,2), (5,1)}
    - B = {ambos dados son impares} = {...}
    - A∩B = {los dados suman 6 y son impares} = {(1,5), (3,3), (5,1)}

  Entonces $P(A) = \frac{5}{36},\ P(B) = \frac{9}{36},\ P(A \cap B) = \frac{3}{36}$, luego:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/36}{9/36} = \frac{1}{3}$$

- Cual es la probabilidad de que los dados sumen 7 dado que ambos son impares?

Probabilidad condicional - Bayes

[Bayes] Si A y B son eventos y la probabilidad de B no es 0, entonces:

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

Probabilidad condicional - Bayes

[Bayes] Si A y B son eventos y la probabilidad de B no es 0, entonces:

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

Ejemplo clásico:

Un test para una enfermedad tiene confiabilidad del 99% (detecta correctamente 99% de casos positivos y negativos). El 1% de la población posee la enfermedad. Si un individuo testea positivo, cuál es la probabilidad de que tenga la enfermedad?

Probabilidad condicional - Bayes

[Bayes] Si A y B son eventos y la probabilidad de B no es 0, entonces:

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

Ejemplo clásico:
Un test para una enfermedad tiene confiabilidad del 99% (detecta correctamente 99% de casos positivos y negativos). El 1% de la población posee la enfermedad. Si un individuo testea positivo, cuál es la probabilidad de que tenga la enfermedad?

B = el test es positivo, A = el individuo tiene la enfermedad.

Probabilidad condicional - Bayes

[Bayes] Si A y B son eventos y la probabilidad de B no es 0, entonces:

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

Ejemplo clásico:
Un test para una enfermedad tiene confiabilidad del 99% (detecta correctamente 99% de casos positivos y negativos). El 1% de la población posee la enfermedad. Si un individuo testea positivo, cuál es la probabilidad de que tenga la enfermedad?

B = el test es positivo, A = el individuo tiene la enfermedad.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Probabilidad condicional - Bayes

[Bayes] Si A y B son eventos y la probabilidad de B no es 0, entonces:

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

Ejemplo clásico:

Un test para una enfermedad tiene confiabilidad del 99% (detecta correctamente 99% de casos positivos y negativos). El 1% de la población posee la enfermedad. Si un individuo testea positivo, cuál es la probabilidad de que tenga la enfermedad?

B = el test es positivo, A = el individuo tiene la enfermedad.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{0.99 \cdot 0.01}{0.0198} = 0.5$$

Variables aleatorias

Hasta ahora definimos S como el universo de cosas que nos interesa medir y trabajamos con eventos.

En la teoría de probabilidad las **variables aleatorias** representan mediciones a valores reales que se toman sobre el espacio de probabilidad.

Variables aleatorias

Hasta ahora definimos *S* como el universo de cosas que nos interesa medir y trabajamos con eventos.

En la teoría de probabilidad las **variables aleatorias** representan mediciones a valores reales que se toman sobre el espacio de probabilidad.

Formalmente, una variable aleatoria *X* es una función

$$X : S \to \mathbb{R}$$

que cumple que para todo intervalo de números reales se puede calcular la probabilidad de $X^{-1}([a, b])$.

Variables aleatorias

Hasta ahora definimos *S* como el universo de cosas que nos interesa medir y trabajamos con eventos.

En la teoría de probabilidad las **variables aleatorias** representan mediciones a valores reales que se toman sobre el espacio de probabilidad.

Formalmente, una variable aleatoria *X* es una función

$$X : S \to \mathbb{R}$$

que cumple que para todo intervalo de números reales se puede calcular la probabilidad de $X^{-1}([a, b])$.

Ejemplos:

- X = la suma de dos dados.
- Y = la temperatura a las 12AM de un dia.
- Z = el peso de un bebé recién nacido.

Variables aleatorias

Hasta ahora definimos *S* como el universo de cosas que nos interesa medir y trabajamos con eventos.

En la teoría de probabilidad las **variables aleatorias** representan mediciones a valores reales que se toman sobre el espacio de probabilidad.

Formalmente, una variable aleatoria *X* es una función

$$X : S \to \mathbb{R}$$

que cumple que para todo intervalo de números reales se puede calcular la probabilidad de $X^{-1}([a, b])$.

Ejemplos:

- X = la suma de dos dados.
- Y = la temperatura a las 12AM de un dia.
- Z = el peso de un bebé recién nacido.

Abuso: observemos que muchas veces cometemos el abuso de no especificar cuál es el espacio *S*.

Variables aleatorias

- Una variable aleatoria se dice **discreta** si toma una cantidad finita* de valores.

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Variables aleatorias

- Una variable aleatoria se dice **discreta** si toma una cantidad finita* de valores.
- Una variable aleatoria se dice **continua** si tiene una función de densidad, es decir una f tal que:

$$P(X \in [a,b]) = \int_a^b f(x)dx$$

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Variables aleatorias

- Una variable aleatoria se dice **discreta** si toma una cantidad finita* de valores.
- Una variable aleatoria se dice **continua** si tiene una función de densidad, es decir una *f* tal que:

$$P(X \in [a,b]) = \int_a^b f(x)dx$$

- En caso contrario se dice **mixta** (pero probablemente no analicemos ningún caso de este estilo en este curso).

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Variables aleatorias

- Una variable aleatoria se dice **discreta** si toma una cantidad finita* de valores.
- Una variable aleatoria se dice **continua** si tiene una función de densidad, es decir una *f* tal que:

$$P(X \in [a, b]) = \int_a^b f(x)dx$$

- En caso contrario se dice **mixta** (pero probablemente no analicemos ningún caso de este estilo en este curso).

En nuestros ejemplos X es discreta, mientras lo más sensato es modelar a Y y Z como variables continuas.

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Variables aleatorias

- Una variable aleatoria se dice **discreta** si toma una cantidad finita* de valores.
- Una variable aleatoria se dice **continua** si tiene una función de densidad, es decir una *f* tal que:

$$P(X \in [a,b]) = \int_a^b f(x)dx$$

- En caso contrario se dice **mixta** (pero probablemente no analicemos ningún caso de este estilo en este curso).

En nuestros ejemplos X es discreta, mientras lo más sensato es modelar a Y y Z como variables continuas.

Definicion: Dada una variable aleatoria *X* definimos su función de distribución acumulada como:

$$F_X(\alpha) = P(X < \alpha)$$

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Vectores aleatorios

Un **vector aleatorio** es una tupla de variables aleatorias $(X_1, X_2, \cdots, X_n)$.

Vectores aleatorios

Un **vector aleatorio** es una tupla de variables aleatorias $(X_1, X_2, \cdots, X_n)$.

Definiciones en 2 variables:

- Un vector aleatorio (X, Y) se dice discreto si toma una cantidad finita* de valores.
- Un vector aleatorio se dice continuo si tiene una densidad, es decir una función tal que:

$$P((X,Y) \in A) = \int_A f_{X,Y}(x,y)\ dxdy$$

para cualquier *caja A*.

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Vectores aleatorios

Un **vector aleatorio** es una tupla de variables aleatorias $(X_1, X_2, \cdots, X_n)$.

Definiciones en 2 variables:

- Un vector aleatorio (X, Y) se dice discreto si toma una cantidad finita* de valores.
- Un vector aleatorio se dice continuo si tiene una densidad, es decir una función tal que:

$$P((X,Y) \in A) = \int_A f_{X,Y}(x,y)\ dxdy$$

para cualquier *caja A*.

Para un vector aleatorio continuo podemos calcular la probabilidad de que (X,Y) se encuentre dentro de una región como la integral de la densidad sobre esa región.

*o numerable. En mayor deatalle, si hay una cantidad finita de valores que toma con probabilidad 1.

Vectores aleatorios

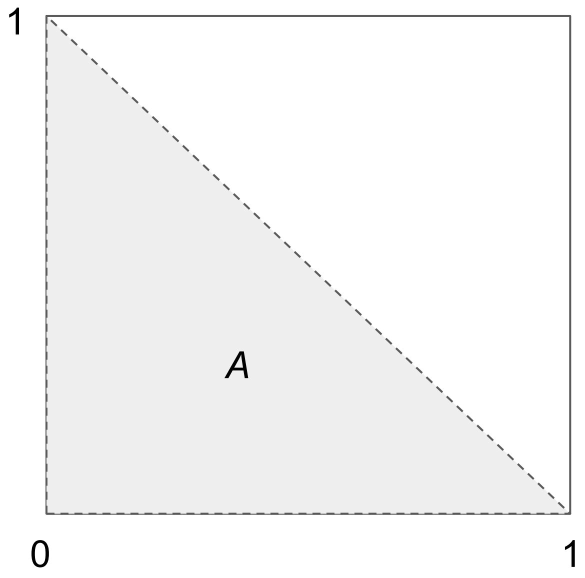
Ejemplo:
Si tomamos dos números x, y al azar en el intervalo [0,1], cuál es la probabilidad de que x+y sea mayor a 1?

Idea:

- Considerar (*X, Y*) un vector aleatorio con densidad *f* igual a 1 en el [0,1]x[0,1] y 0 fuera.

- Queremos calcular:

$$P(X+Y<1) = P((X,Y)\in A) = \int_A 1\ dxdy = \frac{1}{2}$$



Esperanza, Varianza, Covarianza

Un valor importante es la **esperanza** de una variable aleatoria. Es el **valor esperado o promedio** del experimento.

- Para una variable **discreta** definimos:

$$\mathbb{E}(X) = \sum_i P(X = x_i)x_i$$

- Mientras que para una **continua** resulta:

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

Esperanza, Varianza, Covarianza

Algunos ejemplos:

- X = {la suma de 2 dados}

$$\mathbb{E}(X) = P(X=2)\cdot 2 + P(X=3)\cdot 3 + \ldots + P(X=12)\cdot 12 =$$

$$\frac{1}{36}(2+12) + \frac{2}{36}(3+11) + \frac{3}{36}(4+10) + \frac{4}{36}(5+9) + \frac{5}{36}(6+8) + \frac{6}{36}(7) = 7$$

Esperanza, Varianza, Covarianza

Algunos ejemplos:

- X = {la suma de 2 dados}

$$\mathbb{E}(X) = P(X=2)\cdot 2 + P(X=3)\cdot 3 + \ldots + P(X=12)\cdot 12 =$$

$$\frac{1}{36}(2+12) + \frac{2}{36}(3+11) + \frac{3}{36}(4+10) + \frac{4}{36}(5+9) + \frac{5}{36}(6+8) + \frac{6}{36}(7) = 7$$

- W = {el máximo de dos dados}

$$E(W) = 6\cdot\frac{11}{36} + 4\cdot\frac{9}{36} + 4\cdot\frac{7}{36} + 3\cdot\frac{5}{36} + 2\cdot\frac{3}{36} + 1\cdot\frac{1}{36} = 4.472\ldots$$

Esperanza, Varianza, Covarianza

Algunos ejemplos:

- X = {la suma de 2 dados}

$$\mathbb{E}(X) = P(X=2)\cdot 2 + P(X=3)\cdot 3 + \ldots + P(X=12)\cdot 12 =$$

$$\frac{1}{36}(2+12) + \frac{2}{36}(3+11) + \frac{3}{36}(4+10) + \frac{4}{36}(5+9) + \frac{5}{36}(6+8) + \frac{6}{36}(7) = 7$$

- W = {el máximo de dos dados}

$$E(W) = 6\cdot\frac{11}{36} + 4\cdot\frac{9}{36} + 4\cdot\frac{7}{36} + 3\cdot\frac{5}{36} + 2\cdot\frac{3}{36} + 1\cdot\frac{1}{36} = 4.472\ldots$$

- Y = {la cantidad de caras obtenidas al tirar una moneda 4 veces}

Esperanza, Varianza, Covarianza

Algunos ejemplos:

- X = {la suma de 2 dados}

$$\mathbb{E}(X) = P(X=2)\cdot 2 + P(X=3)\cdot 3 + \ldots + P(X=12)\cdot 12 =$$

$$\frac{1}{36}(2+12) + \frac{2}{36}(3+11) + \frac{3}{36}(4+10) + \frac{4}{36}(5+9) + \frac{5}{36}(6+8) + \frac{6}{36}(7) = 7$$

- W = {el máximo de dos dados}

$$E(W) = 6\cdot\frac{11}{36} + 4\cdot\frac{9}{36} + 4\cdot\frac{7}{36} + 3\cdot\frac{5}{36} + 2\cdot\frac{3}{36} + 1\cdot\frac{1}{36} = 4.472\ldots$$

- Y = {la cantidad de caras obtenidas al tirar una moneda 4 veces}
- Z = {la altura de un bebé recién nacido}

Esperanza, Varianza, Covarianza

- La **varianza** y el **desvío estándar** de *X*:

$$Var(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \qquad \sigma(X) = \sqrt{Var(X)}$$

- Si *Y* es otra variable aleatoria en el mismo espacio de probabilidad, la **covarianza** y la **correlación** entre *X* e *Y*:

$$Cov(X,Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) \qquad corr(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Esperanza, Varianza, Covarianza

- La **varianza** y el **desvío estándar** de *X*:

$$Var(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \qquad \sigma(X) = \sqrt{Var(X)}$$

- Si *Y* es otra variable aleatoria en el mismo espacio de probabilidad, la **covarianza** y la **correlación** entre *X* e *Y*:

$$Cov(X,Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) \qquad corr(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Ejemplos:

- X = {la suma de dos dados}, Y = {el doble del valor obtenido al tirar un dado}.

Esperanza, Varianza, Covarianza

- La **varianza** y el **desvío estándar** de *X*:

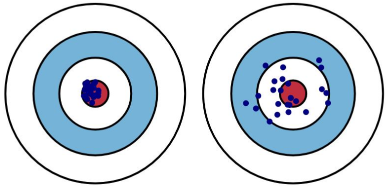
$$Var(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \qquad \sigma(X) = \sqrt{Var(X)}$$

- Si *Y* es otra variable aleatoria en el mismo espacio de probabilidad, la **covarianza** y la **correlación** entre *X* e *Y*:

$$Cov(X,Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) \qquad corr(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Ejemplos:

- X = {la suma de dos dados}, Y = {el doble del valor obtenido al tirar un dado}.



Esperanza, Varianza, Covarianza

Propiedades:

- La esperanza es lineal:

$$\begin{cases} \mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) \\ \mathbb{E}(cX) = c\mathbb{E}(X) \end{cases}$$

- Varianza:

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

$$Var(cX) = c^2 Var(X)$$

- Covarianza:

$$cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$Var(X+Y) = Var(X) + Var(Y) + 2cov(X,Y)$$

Distribuciones

- Dos variables aleatorias tienen la misma distribución, si sus funciones de distribución acumuladas son iguales.

- Esto permite abstraerse definitivamente del espacio muestral y hablar directamente de distribuciones.

Distribuciones

- Dos variables aleatorias tienen la misma distribución, si sus funciones de distribución acumuladas son iguales.

- Esto permite abstraerse definitivamente del espacio muestral y hablar directamente de distribuciones.

Ejemplos:

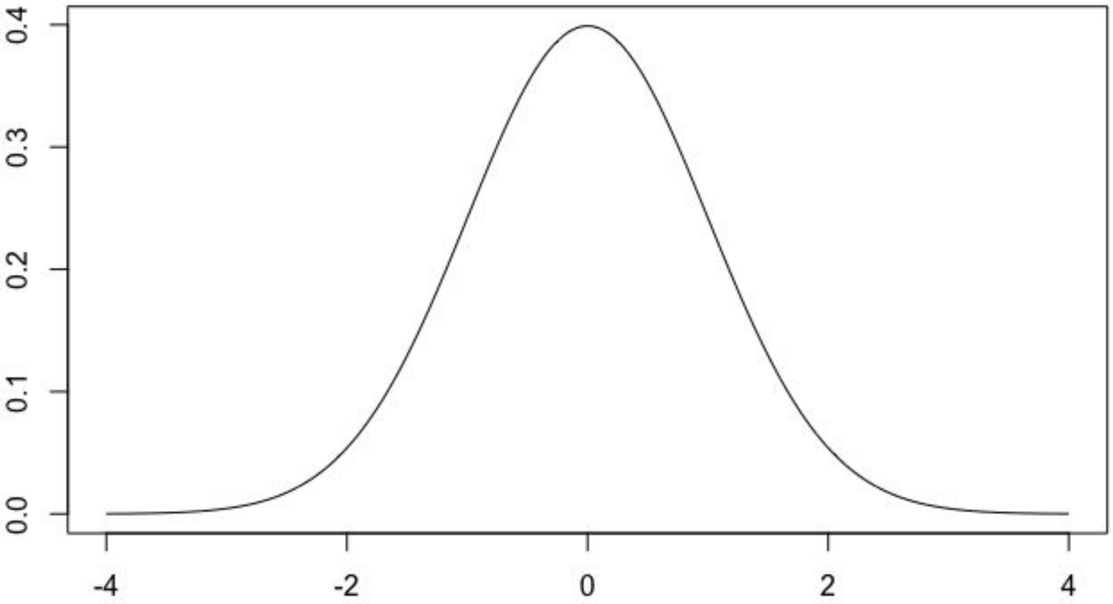
- Una variable aleatoria continua con función de densidad

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

tiene distribución Normal con media 0 y desvío 1. Se denota $X \sim \mathcal{N}(0,1)$.

Distribuciones

Si *X* tiene distribución normal 0,1 entonces $\mathbb{E}(X) = 0$ y $Var(X) = 1$.



Distribuciones

Otros ejemplos notables:

- [Bernoulli] $Ber(p)$: evento binario. Vale 1 con probabilidad p y 0 con probabilidad 1-p.
- [Binomial] $Bin(n, p)$: cantidad de resultados positivos en n repeticiones de un evento $Ber(p)$.
- [Geometrica] $\mathcal{G}(p)$: cantidad de repeticiones de un evento $Ber(p)$ hasta obtener el primer positivo.
- [Poisson] $Poisson(\lambda)$: cantidad de ocurrencias de un evento de Poisson en un intervalo de tiempo.

Distribuciones

Otros ejemplos notables:

- [Exponencial] $exp(\lambda)$: tiempo hasta la primera ocurrencia de un evento de Poisson.
- [Uniforme] $\mathcal{U}([a, b])$: distribución que le da la misma probabilidad a todos los puntos de un segmento.
- [Normal] $\mathcal{N}(\mu, \sigma)$: variable igual a $\mu + \sigma Z$ con $Z \sim \mathcal{N}(0, 1)$.
- [Lognormal] $Lognormal(\mu, \sigma)$: variable aleatoria cuyo logaritmo es $\mathcal{N}(\mu, \sigma)$.
- [Normal multivariada] $\mathcal{N}(\mu, \Sigma)$: es un vector aleatorio continuo con función de densidad:

$$f_X(x) = (2\pi)^{-\frac{k}{2}} \det(\Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^t \Sigma (x-\mu)}$$

Lognormal como modelo de asset pricing

Supongamos un precio de un asset *S*. Nos interesa entender su retorno relativo:

$$\frac{S(t+\delta t) - S(t)}{\delta t S(t)}$$

Lognormal como modelo de asset pricing

Supongamos un precio de un asset *S*. Nos interesa entender su retorno relativo:

$$\frac{S(t+\delta t)-S(t)}{\delta t S(t)}$$

Si asumimos [mal!] que este asset tiene un retorno constante en el tiempo tendríamos:

$$\frac{S(t+\delta t)-S(t)}{\delta t S(t)} = \mu \longrightarrow S'(t) = \mu S(t)$$

Ya conocemos una ecuación parecida, que da como resultado $S(t) = S(0)e^{\mu t}$.

Lognormal como modelo de asset pricing

Supongamos un precio de un asset *S*. Nos interesa entender su retorno relativo:

$$\frac{S(t+\delta t)-S(t)}{\delta t S(t)}$$

Si asumimos [mal!] que este asset tiene un retorno constante en el tiempo tendríamos:

$$\frac{S(t+\delta t)-S(t)}{\delta t S(t)} = \mu \longrightarrow S'(t) = \mu S(t)$$

Ya conocemos una ecuación parecida, que da como resultado $S(t) = S(0)e^{\mu t}$.

Una asunción más razonable es:

$$\frac{S(t+\delta t)-S(t)}{\delta t S(t)} = \mu + \sigma \frac{1}{\sqrt{\delta t}} Z \sim N(\mu, \tilde{\sigma})$$

Análogamente a la ecuación de retorno fijo esto implica que S tiene distribución lognormal.

Independencia

Dos eventos A y B se dicen independientes si

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

Esto equivale a

$$P(A|B) = P(A) \text{ y } P(A \cap B) = P(A)P(B)$$

Dos eventos son independientes cuando saber que uno sucedió no afecta la probabilidad de que el otro suceda.

Independencia

Dos eventos A y B se dicen independientes si

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

Esto equivale a

$$P(A|B) = P(A) \text{ y } P(A \cap B) = P(A)P(B)$$

Dos eventos son independientes cuando saber que uno sucedió no afecta la probabilidad de que el otro suceda.

En nuestro ejemplo de los dos dados:

- A = {el primer dado es igual a 1} y B = {el segundo dado es igual a 1} son eventos independientes.
- A = {la suma es par} y B = {la suma es impar} **no** son eventos independientes.
- A = {la suma es mayor a 9} y B = {el primer dado es par} **no** son eventos independientes.

Independencia

Dos variables aleatorias *X*, *Y* son independientes si $X^{-1}([a,b])$ y $Y^{-1}([c,d])$ son eventos independientes para cualquier par de intervalos [*a*,*b*] y [*c*,*d*].

Independencia

Dos variables aleatorias *X*, *Y* son independientes si $X^{-1}([a,b])$ y $Y^{-1}([c,d])$ son eventos independientes para cualquier par de intervalos [*a*,*b*] y [*c*,*d*].

Intuición: dos experimentos son independientes si cualquier medición que conozcamos sobre uno de ellos no afecta la probabilidad de obtener otra medición sobre el otro.

Independencia

Dos variables aleatorias *X*, *Y* son independientes si $X^{-1}([a,b])$ y $Y^{-1}([c,d])$ son eventos independientes para cualquier par de intervalos [*a*,*b*] y [*c*,*d*].

Intuición: dos experimentos son independientes si cualquier medición que conozcamos sobre uno de ellos no afecta la probabilidad de obtener otra medición sobre el otro.

Importante: dos variables aleatorias independientes siempre van a tener covarianza igual a 0. El recíproco **no es cierto**.

Independencia

Ejemplos:

- En el ejemplo en el que lanzamos dos dados:
  - La suma de los dados y el máximo de los dados **NO** son independientes.
  - El valor del primer dado y el valor del segundo dado **SI** son independientes.

Independencia

Ejemplos:

- En el ejemplo en el que lanzamos dos dados:
  - La suma de los dados y el máximo de los dados **NO** son independientes.
  - El valor del primer dado y el valor del segundo dado **SI** son independientes.
- El peso y la altura de un recién nacido deberían ser modelados como variables aleatorias **NO** independientes.

Independencia

Si *X* e *Y* son variables aleatorias independientes:

- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$
- $Var(X + Y) = Var(X) + Var(Y)$
- $cov(X, Y) = 0$

Independencia

Si *X* e *Y* son variables aleatorias independientes:

- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$
- $Var(X+Y) = Var(X) + Var(Y)$
- $cov(X,Y) = 0$

Si *X* e *Y* son continuas e independientes entonces:

- $f_{XY}(x,y) = f_X(x)f_Y(y)$
- $f_{X+Y}(x) = (f_X * f_Y)(x) = \int_{-\infty}^{\infty} f_X(z)f_Y(x-z)dz$

Una consecuencia importante:

Si *X* e *Y* son dos variables aleatorias normales e independientes entonces *X+Y* también tiene distribución normal.

La ley de los grandes números

[Ley de los grandes números] Dada una sucesión de variables aleatorias X_i, todas independientes, con esperanza finita y con la misma distribución (iid) vale que:

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \rightarrow \mathbb{E}(X_1)$$

cuando *n* tiende a infinito.

La ley de los grandes números

[Ley de los grandes números] Dada una sucesión de variables aleatorias X_i, todas independientes, con esperanza finita y con la misma distribución (iid) vale que:

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \to \mathbb{E}(X_1)$$

cuando *n* tiende a infinito.

Intuicion

- La esperanza se puede interpretar como el "promedio de todos los posibles valores" que toma una variable aleatoria.
- La suma del lado izquierdo se puede interpretar como un promedio de infinitos valores de la misma variable aleatoria.

La ley de los grandes números

Ejemplos.

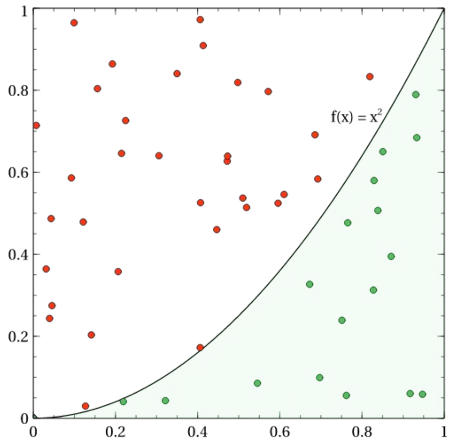
- Si lanzamos una moneda infinitas veces, la proporción de caras que obtengamos va a converger a 0.5.

La ley de los grandes números

Ejemplos.

- Si lanzamos una moneda infinitas veces, la proporción de caras que obtengamos va a converger a 0.5.

- Elegimos un punto al azar en [0,1]x[0,1] y sumamos 1 si cae debajo de la curva $y = x^2$ y 0 en caso contrario.
  La proporción de 1s obtenida converge a la integral: $\int_0^1 x^2 dx$



La ley de los grandes números

"Demostración": llamemos A a la región debajo de la curva $y = x^2$.

- Si $U \sim \mathcal{U}([0,1] \times [0,1])$ entonces la proporción de unos tiene que converger a la esperanza de la variable aleatoria V que vale 1 si $U \in A$ y 0 si no.

La ley de los grandes números

"Demostración": llamemos A a la región debajo de la curva $y = x^2$.

- Si $U \sim \mathcal{U}([0,1] \times [0,1])$ entonces la proporción de unos tiene que converger a la esperanza de la variable aleatoria V que vale 1 si $U \in A$ y 0 si no.

- Vale que

$$\mathbb{E}(V) = 0 \cdot P(U \notin A) + 1 \cdot P(U \in A) = P(U \in A)$$

La ley de los grandes números

"Demostración": llamemos A a la región debajo de la curva $y = x^2$.

- Si $U \sim \mathcal{U}([0,1] \times [0,1])$ entonces la proporción de unos tiene que converger a la esperanza de la variable aleatoria V que vale 1 si $U \in A$ y 0 si no.

- Vale que

$$\mathbb{E}(V) = 0 \cdot P(U \notin A) + 1 \cdot P(U \in A) = P(U \in A)$$

- Y ademas:

$$P(U \in A) = \int_A f_U(x,y)dxdy = \int_A 1 \; dxdy = \int_0^1 x^2 dx$$

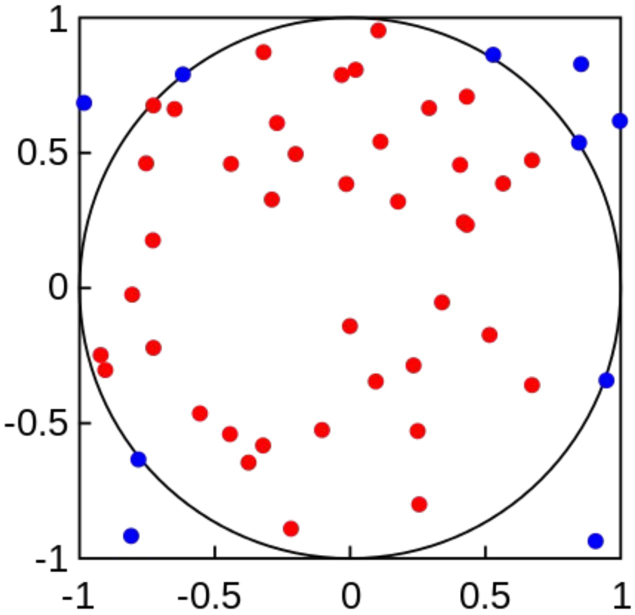
La ley de los grandes números

- Esta idea se llama integración de Monte Carlo.

- En general, los métodos que utilizan simulaciones para aproximar una esperanza se llaman Métodos de Monte Carlo.

- Así como existen ecuaciones diferenciales que modelan precios de derivados, veremos que estos también se pueden interpretar como esperanzas de variables aleatorias.

- Los métodos Monte Carlo son muy versátiles, pero su convergencia suele ser *lenta*.



Teorema Central del Límite

[Teorema Central del Límite] Dada una sucesión de variables aleatorias X_i todas independientes, con esperanza μ y varianza σ^2 y con la misma distribución (iid) llamemos

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Entonces cuando *n* tiende a infinito la distribución de

$$Z_n = \sqrt{n} \cdot \frac{\overline{X}_n - \mu}{\sigma}$$

Converge a una $\mathcal{N}(0, 1)$.

Teorema Central del Límite

[Teorema Central del Límite] Dada una sucesión de variables aleatorias X_i todas independientes, con esperanza μ y varianza σ^2 y con la misma distribución (iid) llamemos

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Entonces cuando *n* tiende a infinito la distribución de

$$Z_n = \sqrt{n} \cdot \frac{\overline{X}_n - \mu}{\sigma}$$

Converge a una $\mathcal{N}(0, 1)$.

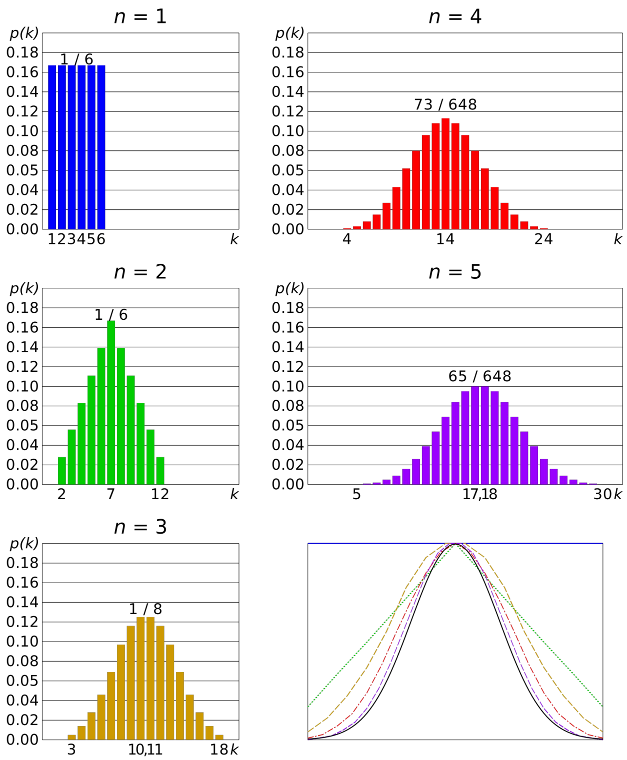
Intuición/Consecuencias

- No hay. Es un teorema medio mágico.
- La velocidad de convergencia de la Ley de los Grandes Números es $O\left(\frac{1}{\sqrt{n}}\right)$.

Teorema Central del Límite

Ejemplos

- Si tiramos n dados (para n grande) y calculamos la suma de los valores que obtuvimos entonces la distribución es aproximadamente normal.

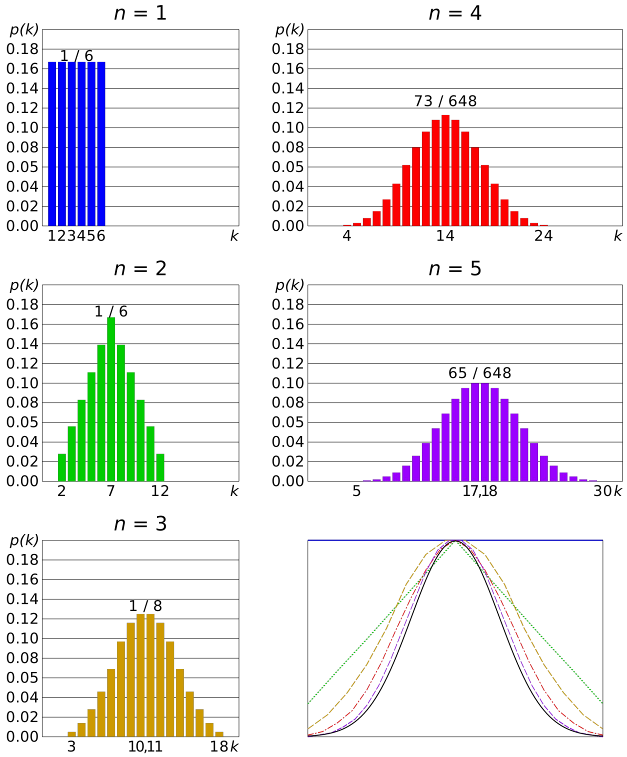


Teorema Central del Límite

Ejemplos

- Si tiramos n dados (para n grande) y calculamos la suma de los valores que obtuvimos entonces la distribución es aproximadamente normal.

- Si tiramos una moneda 30 veces, cual es la probabilidad de obtener entre 10 y 20 caras? (aproximadamente)



Teorema Central del Límite

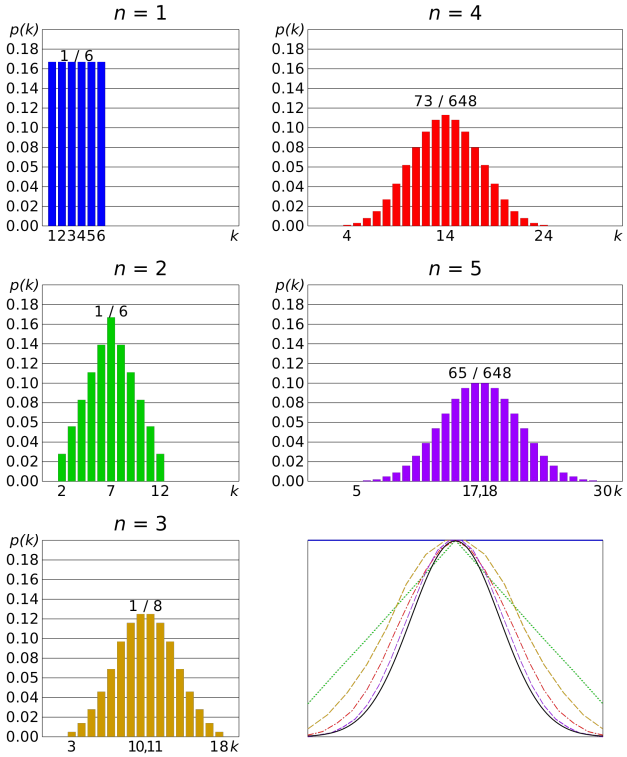
Ejemplos

- Si tiramos n dados (para n grande) y calculamos la suma de los valores que obtuvimos entonces la distribución es aproximadamente normal.

- Si tiramos una moneda 30 veces, cual es la probabilidad de obtener entre 10 y 20 caras? (aproximadamente)

  Queremos saber $P(10 \leq X \leq 20)$ para $X = Bin(30, 0.5)$. Observamos que

$$X = B_1 + \ldots + B_{30}$$



Teorema Central del Límite

Ejemplos

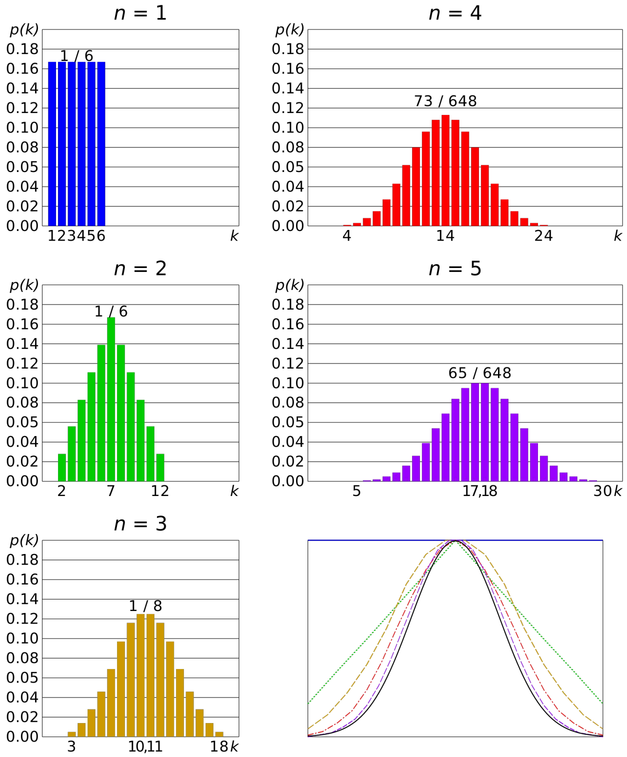
- Si tiramos n dados (para n grande) y calculamos la suma de los valores que obtuvimos entonces la distribución es aproximadamente normal.

- Si tiramos una moneda 30 veces, cual es la probabilidad de obtener entre 10 y 20 caras? (aproximadamente)

  Queremos saber $P(10 \leq X \leq 20)$ para $X = Bin(30, 0.5)$. Observamos que

$$X = B_1 + \ldots + B_{30}$$

  Luego $Z = \sqrt{30} \cdot \dfrac{X/30 - 0.5}{0.25}$ tiene una distribución casi $\mathcal{N}(0, 1)$.



Teorema Central del Límite

En Machine Learning

- Tests estadisticos:
  El TCL juega un papel fundamental en la mayoría de los tests estadísticos. Ejemplo más simple: Z-test.

- Estimaciones de errores:
  Muchas estimaciones de parámetros se basan en tomar promedios. El TCL nos permite construir intervalos de error para esas estimaciones.

- Asunciones en modelos:
  Muchos modelos asumen normalidad en distintas cantidades (ejemplo: una regresión lineal). En general esto se encuentra fuertemente influenciado por el TCL.

Esperanza condicional

Dadas *X, Y* variables aleatorias discretas definimos:

$$\mathbb{E}(X|Y=y) = \sum_x x P(X=x|Y=y) = \sum_x x \frac{p(x,y)}{p_Y(y)}$$

Es el valor esperado de X dada una observación de Y.

Esperanza condicional

Dadas *X, Y* variables aleatorias discretas definimos:

$$\mathbb{E}(X|Y = y) = \sum_x xP(X = x|Y = y) = \sum_x x\frac{p(x, y)}{p_Y(y)}$$

Es el valor esperado de X dada una observación de Y.

Ejemplo:

Si X e Y son Bernoulli, cual es el valor esperado de X si sabemos que X+Y=1?

Esperanza condicional

Dadas *X, Y* variables aleatorias discretas definimos:

$$\mathbb{E}(X|Y=y) = \sum_x xP(X=x|Y=y) = \sum_x x\frac{p(x,y)}{p_Y(y)}$$

Es el valor esperado de X dada una observación de Y.

Ejemplo:
Si X e Y son Bernoulli, cual es el valor esperado de X si sabemos que X+Y=1?

$$\mathbb{E}(X|X+Y=1)$$

Esperanza condicional

Dadas *X, Y* variables aleatorias discretas definimos:

$$\mathbb{E}(X|Y=y) = \sum_x xP(X=x|Y=y) = \sum_x x\frac{p(x,y)}{p_Y(y)}$$

Es el valor esperado de X dada una observación de Y.

Ejemplo:
Si X e Y son Bernoulli, cual es el valor esperado de X si sabemos que X+Y=1?

$$\mathbb{E}(X|X+Y=1) = 0P(X=0|X+Y=1) + 1P(X=1|X+Y=1)$$

Esperanza condicional

Dadas *X, Y* variables aleatorias discretas definimos:

$$\mathbb{E}(X|Y=y) = \sum_x xP(X=x|Y=y) = \sum_x x\frac{p(x,y)}{p_Y(y)}$$

Es el valor esperado de X dada una observación de Y.

Ejemplo:
Si X e Y son Bernoulli, cual es el valor esperado de X si sabemos que X+Y=1?

$$\mathbb{E}(X|X+Y=1) = 0P(X=0|X+Y=1) + 1P(X=1|X+Y=1) = \frac{P(X=1, X+Y=1)}{P(X+Y=1)}$$

Esperanza condicional

Dadas *X, Y* variables aleatorias discretas definimos:

$$\mathbb{E}(X|Y=y) = \sum_x xP(X=x|Y=y) = \sum_x x\frac{p(x,y)}{p_Y(y)}$$

Es el valor esperado de X dada una observación de Y.

Ejemplo:

Si X e Y son Bernoulli, cual es el valor esperado de X si sabemos que X+Y=1?

$$\mathbb{E}(X|X+Y=1) = 0P(X=0|X+Y=1) + 1P(X=1|X+Y=1) = \frac{P(X=1, X+Y=1)}{P(X+Y=1)} = \frac{1}{2}$$

Esperanza condicional

De la misma forma si *X, Y* son variables aleatorias continuas definimos la densidad condicional:

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

y la esperanza condicional:

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x) dx$$

Esperanza condicional

La **esperanza condicional** es la variable aleatoria $\mathbb{E}(X|Y)$ definida como

$$\mathbb{E}(X|Y)(y) = \mathbb{E}(X|Y = y)$$

Esperanza condicional

La **esperanza condicional** es la variable aleatoria $\mathbb{E}(X|Y)$ definida como

$$\mathbb{E}(X|Y)(y) = \mathbb{E}(X|Y = y)$$

La esperanza condicional es nuestro mejor estimado de *X* si conociéramos a *Y*, es de algún modo la variable que canaliza la regla para convertir un valor de *Y* en un estimado de *X*.

Esperanza condicional

La **esperanza condicional** es la variable aleatoria $\mathbb{E}(X|Y)$ definida como

$$\mathbb{E}(X|Y)(y) = \mathbb{E}(X|Y = y)$$

La esperanza condicional es nuestro mejor estimado de *X* si conociéramos a *Y*, es de algún modo la variable que canaliza la regla para convertir un valor de *Y* en un estimado de *X*.

Propiedades importantes:

- $\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$
- Lo que para variables discretas es $\mathbb{E}(X) = \sum_y \mathbb{E}(X|Y = y)P(Y = y)$
- Y para variables continuas es $\mathbb{E}(X) = \int_{-\infty}^{+\infty} \mathbb{E}(X|Y = y) f_Y(y) dy$

Esperanza condicional

Ejemplo: un trabajador está encerrado en una mina con 3 puertas.

- La primera puerta lleva a la salida en 3 horas.
- La segunda puerta lleva de vuelta a la mina en 5 horas.
- La tercera puerta lleva de vuelta a la mina en 7 horas.

Cual es el tiempo que en promedio debería llevar al trabajador salir de la mina?

Esperanza condicional

Ejemplo: un trabajador está encerrado en una mina con 3 puertas.

- La primera puerta lleva a la salida en 3 horas.
- La segunda puerta lleva de vuelta a la mina en 5 horas.
- La tercera puerta lleva de vuelta a la mina en 7 horas.

Cual es el tiempo que en promedio debería llevar al trabajador salir de la mina?

Si *X* = tiempo en salir de la mina e *Y* = puerta elegida inicialmente, entonces:

$$\mathbb{E}(X) = \mathbb{E}(X|Y=1)P(Y=1) + \mathbb{E}(X|Y=2)P(Y=2) + \mathbb{E}(X|Y=3)P(Y=3)$$

Esperanza condicional

Ejemplo: un trabajador está encerrado en una mina con 3 puertas.

- La primera puerta lleva a la salida en 3 horas.
- La segunda puerta lleva de vuelta a la mina en 5 horas.
- La tercera puerta lleva de vuelta a la mina en 7 horas.

Cual es el tiempo que en promedio debería llevar al trabajador salir de la mina?

Si X = tiempo en salir de la mina e Y = puerta elegida inicialmente, entonces:

$$\mathbb{E}(X) = \mathbb{E}(X|Y=1)P(Y=1) + \mathbb{E}(X|Y=2)P(Y=2) + \mathbb{E}(X|Y=3)P(Y=3)$$

Pero ademas:

- $\mathbb{E}(X|Y=1) = 3$
- $\mathbb{E}(X|Y=2) = 5 + \mathbb{E}(X)$
- $\mathbb{E}(X|Y=3) = 7 + \mathbb{E}(X)$

Esperanza condicional

Ejemplo: un trabajador está encerrado en una mina con 3 puertas.

- La primera puerta lleva a la salida en 3 horas.
- La segunda puerta lleva de vuelta a la mina en 5 horas.
- La tercera puerta lleva de vuelta a la mina en 7 horas.

Cual es el tiempo que en promedio debería llevar al trabajador salir de la mina?

Si *X* = tiempo en salir de la mina e *Y* = puerta elegida inicialmente, entonces:

$$\mathbb{E}(X) = \mathbb{E}(X|Y=1)P(Y=1) + \mathbb{E}(X|Y=2)P(Y=2) + \mathbb{E}(X|Y=3)P(Y=3)$$

Pero ademas:

- $\mathbb{E}(X|Y=1) = 3$
- $\mathbb{E}(X|Y=2) = 5 + \mathbb{E}(X)$
- $\mathbb{E}(X|Y=3) = 7 + \mathbb{E}(X)$

Lo que nos permite despejar $\mathbb{E}(X) = 15$