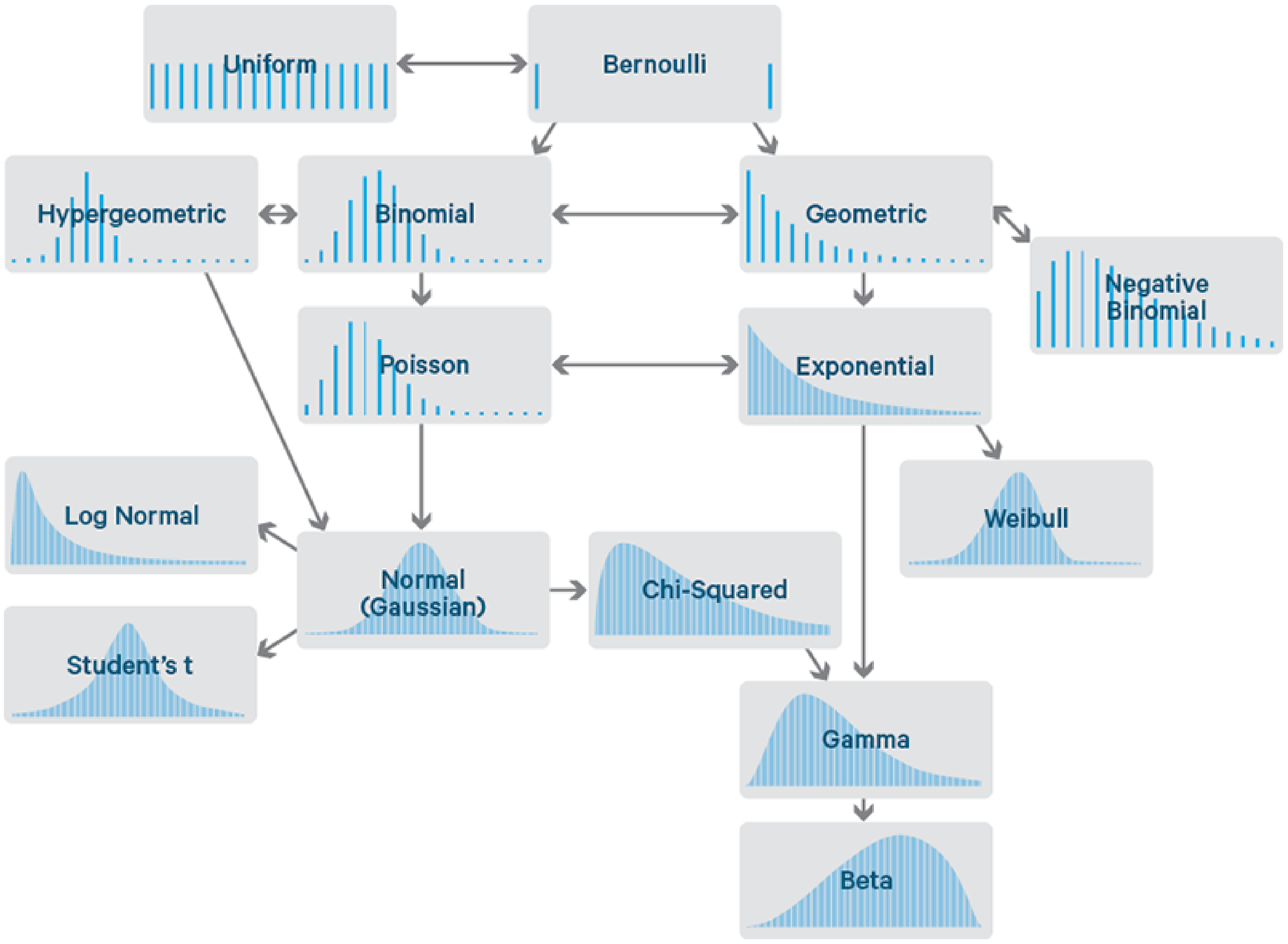
ESTADÍSTICA Y HERRAMIENTAS COMPUTACIONALES

Presentar elementos estadísticos, probabilísticos y herramientas necesarias para soportar decisiones financieras.

Distribuciones

- Tipos de distribuciones y sus derivaciones



Distribuciones

- <u>**Discretas**</u>: Son las cuales la variable de análisis toma solamente algunos valores respecto al estudio realizado.

- <u>**Continuas**</u>: La variable que se estudia puede tomar cualquier valor de un intervalo establecido o definido en el estudio.

Distribuciones Discretas

Distribución de probabilidad binomial

- Características:
  - Número fijo n de pruebas
  - Resultados S o F (Éxito or Falla)
  - Probabilidad de éxito p es el mismo de una prueba a otra. Probabilidad de fracaso q=(1-p)
  - Pruebas son independientes
  - Variable de interés Y, es el número de éxitos durante las n pruebas

Distribución de probabilidad binomial

- Definición: La variable Y tiene una distribución de este tipo basada en sus n pruebas con probabilidad p de éxito si y sólo si →

$$p(y) = \binom{n}{y} p^y q^{n-y} \quad \mathrm{y} = 0,1,2,\dots,\mathrm{n} \quad 0 \leq p \leq 1$$

$$\mu = E(Y) = np \qquad \sigma^2 = V(Y) = npq$$

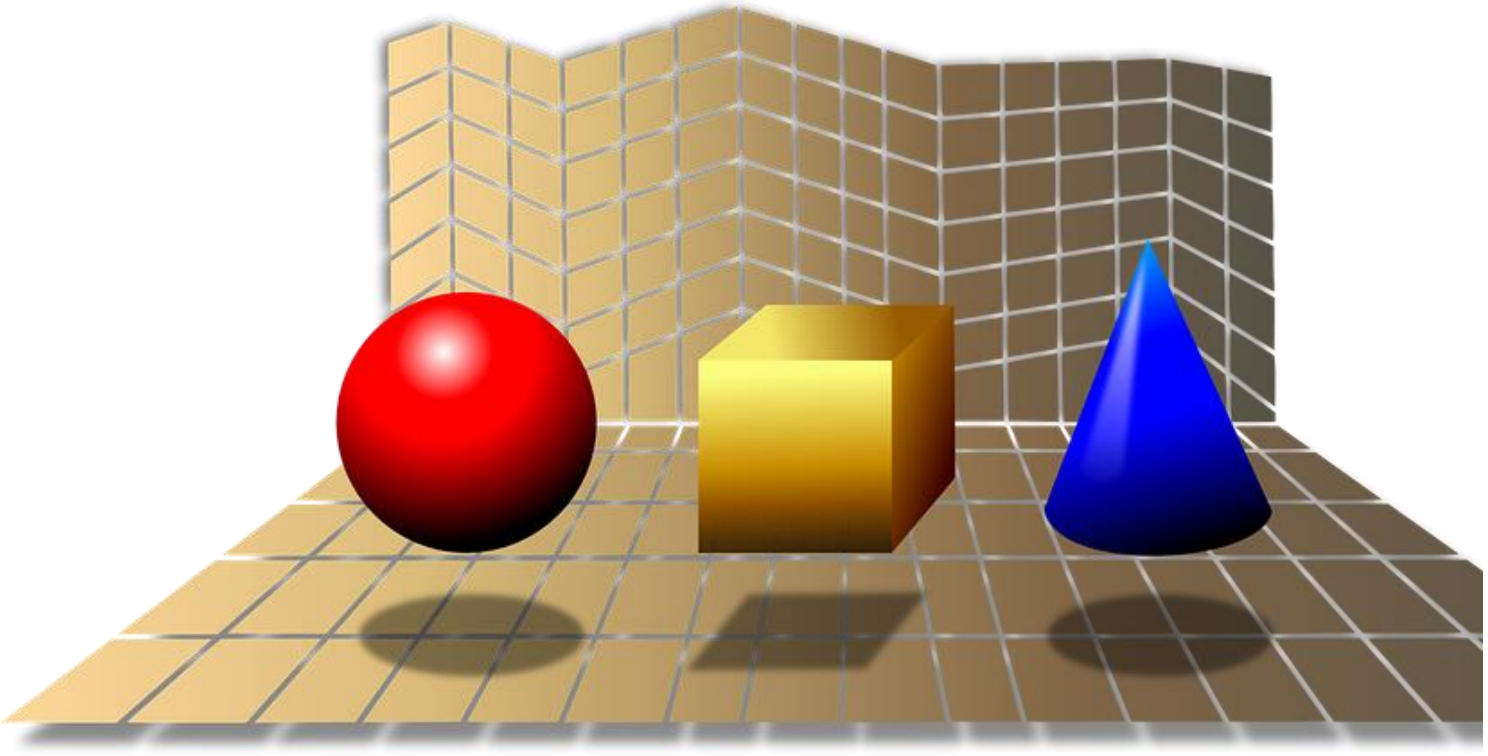
Ejemplo…

- Un trader tiene probabilidad de ganar del 55%, calcular la probabilidad que al hacer 6 negociaciones este trader gane en: a) Cuatro negociaciones, b) Todas las negociaciones y c) Ninguna negociación

- B(n,p) → B(6,55%), q = (1-55%) = 45%
- $p(y = 4) = \binom{6}{4} 55\%^{4} 45\%^{6-4} = 27{,}79\%$
- $p(y = 6) = \binom{6}{6} 55\%^{6} 45\%^{6-6} =?$
- $p(y =?\,) = \binom{6}{?} 55\%^{?} 45\%^{6-?} =?$

Aplicación Python…

Distribuciones Geométrica



Distribución de probabilidad Geométrica

- Características:
  - Similares a una prueba binomial, la única diferencia es que concluye en el primer éxito del suceso evaluado.

Distribución de probabilidad geométrica

- Definición: La variable Y tiene una distribución de este tipo si y sólo si →

$$p(y) = p \quad q^{y-1} \quad \mathrm{y} = 0,1,2,\ldots,\mathrm{n} \;\; 0 \leq p \leq 1$$

$$\mu = E(Y) = \frac{1}{p} \qquad \sigma^2 = V(Y) = \frac{1-p}{p^2}$$

Ejemplo…

- La probabilidad de fallar un pronostico de movimientos de una acción durante cualquier periodo de tiempo es de p=2% y Y denota el número de intervalos de una hora hasta el primer fallo del pronóstico, cuál es la media y la desviación estándar de Y.

- $E(Y) = \frac{1}{p} = \frac{1}{2\%} = 50$: Es decir esperaríamos unas pocas horas antes de encontrar un fallo
- $\mathrm{V}(Y) = \frac{98\%}{0{,}0004} = 2450, \sigma = 49{,}497$

Aplicación Python…

Distribuciones Hipergeométrica

Distribución de probabilidad hipergeométrica

- Características:
  - De una muestra se seleccionan varios elementos al azar con cierto atributo (sin reemplazo).
  - Se desea hallar la probabilidad que los elementos seleccionados provienen del grupo con el atributo definido por el analista.

Distribución de probabilidad hipergeométrica

- Definición: La variable Y tiene una distribución de este tipo si y sólo si →

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \quad \mathrm{y} = 0,1,2, \ldots, \mathrm{n}, \mathrm{y} \leq r, \mathrm{n} - \mathrm{y} \leq N - r$$

$$\mu = E(Y) = \frac{nr}{N} \qquad \sigma^2 = V(Y) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Ejemplo…

- Una compañía de bolsa tiene definido que un comercial atiende a 20 clientes. Es muy costoso realizar pruebas para determinar si una asesoría fue mal realizada, entonces, en lugar de validar el 100% de los comerciales, se diseña un plan de muestreo, con el fin de minimizar el número de asesorías mal realizadas. Donde se pretende validar al menos 5 asesorías de cada comercial y despedir el comercial en caso de mas de una asesoría mal dada. Si un comercial realiza cuatro asesorías malas. ¿Cuál es la probabilidad que sea despedido? ¿Cuál es el número esperado de malas asesorías en la muestra de 5? ¿Cuál es la varianza del número de malas asesorías de la muestra de 5?

- Denotemos a E como el número de asesorías malas de la muestra.
- N=20, r=4 y n=5, el comercial será despedido si Y=2,3 o 4

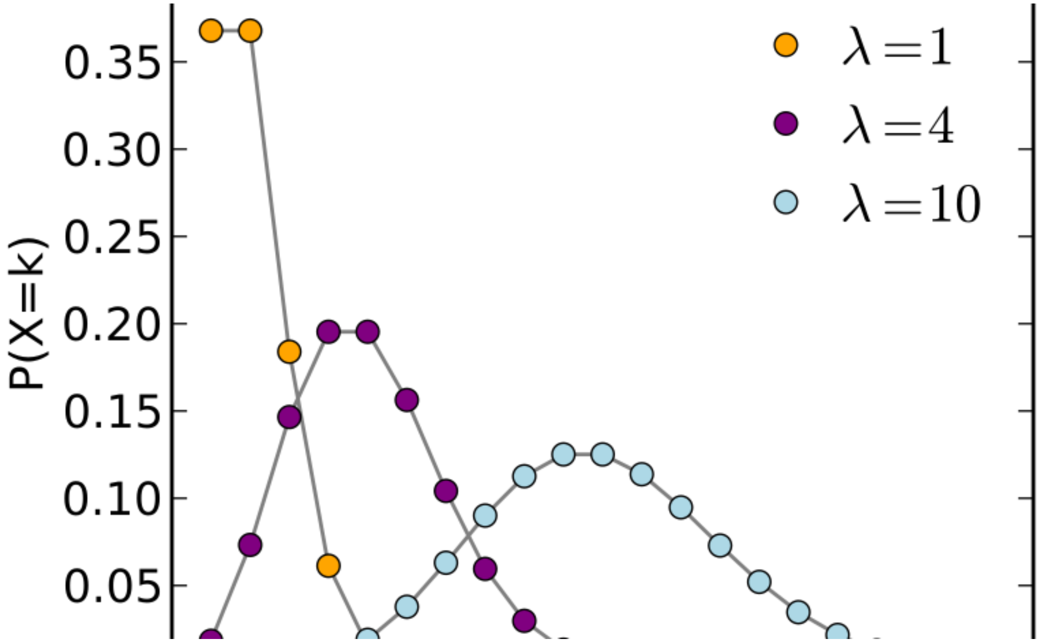
Ejemplo continuación…

- $P(despedir\ ccial) = P(Y \geq 2) = p(2) + p(3) + p(4)$
- $= 1 - p(0) - p(1)$
- $= 1 - \frac{\binom{4}{0}\binom{16}{5}}{\binom{20}{5}} - \frac{\binom{4}{1}\binom{16}{4}}{\binom{20}{5}} = 1 - 0{,}2817 - 0{,}4696 = 24{,}87\%$

- $\mu = \frac{5\times 4}{20} = 1 \qquad y \qquad \sigma^2 = 5\left(\frac{4}{20}\right)\left(\frac{20-4}{20}\right)\left(\frac{20-5}{20-1}\right) = 63{,}2\%$

Aplicación Python…

Distribuciones Poisson



Distribución de probabilidad Poisson

- Características:
- Describir la cantidad de "y" eventos en un lapso de tiempo establecido.
- Si los eventos son independientes y se producen a una velocidad µ
- UNA DE LAS MÁS USADAS

Distribución de probabilidad Poisson

- Definición: La variable Y tiene una distribución de este tipo si y sólo si →

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda} = \frac{\mu^y}{y!} e^{-\mu} \mathrm{y} = 0,1,2,\ldots,\mathrm{n}, \lambda > 0$$

$$\mu = E(Y) = \lambda \qquad \sigma^2 = V(Y) = \lambda$$

Ejemplo…

- Una empresa se equivoca registrando operaciones de acuerdo con un proceso Poisson con un promedio de 3 errores por mes. Durante los últimos dos meses se presentaron 10 errores operativos en registro. ¿Este número es atípico?, ¿De acuerdo con su media es todavía igual a 3?, ¿Indica un aumento en el promedio de incidentes por mes?

- En dos meses (2), la media sería de acuerdo con su distribución
- $\lambda = 2 \times 3 = 6, entonces$ …
- $P(Y \geq 10) = \sum_{y=10}^{\infty} \frac{6^y e^{-6}}{y!} =$ $Opción\ 1\ tabla\ de\ probabilidades$=Opción 2 Python: 4,26%

Continuación

- Con la regla empírica podemos tener que:
  - $\mu = \lambda = 6, \ \sigma^2 = \lambda = 6, \quad \sigma = \sqrt{6} = 2,45$
  - Esta regla nos dice que deberíamos esperar que Y tome valores entre:
  - $\mu \pm 2\sigma = 6 + 2 \times 2,45 = 10,90$
  - En Python que sea menor a 10 es de 95,73% y que sea mayor es de 4,26%
  - Si el número observado es 10 no esta dentro del rango, pero esta cerca por lo que no es improbable que pase, pero puede ser suficientemente improbable para garantizar una investigación.

Aplicación Python…

Distribuciones Continuas

Distribución de probabilidad Uniforme

- Características:
  - Todos los valores tienen una probabilidad muy similar.

Distribución de probabilidad Uniforme

- Definición: La variable Y tiene una distribución de este tipo si y sólo si →

$$f(y) = \begin{cases} \dfrac{1}{\Theta_2 - \Theta_1}, \Theta_2 \leq y \leq \Theta_1 \\ 0, \qquad en\ otro\ punto \end{cases}$$

$$\mu = E(Y) = \frac{\Theta_1 + \Theta_2}{2}$$

$$\sigma^2 = V(Y) = \frac{(\Theta_2 - \Theta_1)^2}{12}$$

Ejemplo

- La llegada de clientes a una caja en un establecimiento sigue una distribución de Uniforme. Se sabe que durante un periodo determinado de 30 minutos, un cliente llega a la caja. Encuentre la probabilidad de que el cliente llegue durante los últimos 5 minutos del periodo de 30 minutos.

- $P(25 \leq Y \leq 30) = \int_{25}^{30} \frac{1}{30} dy = \frac{30-25}{30} = 1/6$

Distribución de probabilidad normal

- Características:
  - También llamada Distr. Gauss
  - La más usada
  - Describe fenómenos naturales fácilmente
  - Errores de ciertas magnitudes

Distribución de probabilidad Normal

- Definición: La variable Y tiene una distribución de este tipo si y sólo si → σ>0 y -∞<μ<∞

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)/(2\sigma^2)} \qquad -\infty \leq y \leq \infty$$

$$\mu = E(Y) \qquad\qquad \sigma^2 = V(Y)$$

Distribución de probabilidad gamma

- Características:
  - Para variables no negativas
  - Distribuciones sesgadas (no simetricas) a la derecha

Distribución de probabilidad Gamma

DEFINICIÓN 4.9 Se dice que una variable aleatoria Y tiene una *distribución gamma con parámetros* $\alpha > 0$ y $\beta > 0$ si y sólo si la función de densidad de Y es

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & 0 \le y < \infty, \\ 0, & \text{en cualquier otro punto,} \end{cases}$$

donde

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\,dy.$$

TEOREMA 4.8 Si Y tiene una distribución gamma con parámetros α y β, entonces

$$\mu = E(Y) = \alpha\beta \quad \text{y} \quad \sigma^2 = V(Y) = \alpha\beta^2.$$

Distr. Gamma

Casos especiales de Gamma

DEFINICIÓN 4.10

Sea ν un entero positivo. Se dice que una variable aleatoria Y tiene *distribución ji cuadrada con* ν *grados de libertad* si y sólo si Y es una variable aleatoria con distribución gamma y parámetros $\alpha = \nu/2$ y $\beta = 2$.

TEOREMA 4.9

Si Y es una variable aleatoria ji cuadrada con ν grados de libertad, entonces

$$\mu = E(Y) = \nu \quad \text{y} \quad \sigma^2 = V(Y) = 2\nu.$$

DEFINICIÓN 4.11

Se dice que una variable aleatoria Y tiene una *distribución exponencial con parámetro* $\beta > 0$ si y sólo si la función de densidad de Y es

$$f(y) = \begin{cases} \dfrac{1}{\beta} e^{-y/\beta}, & 0 \le y < \infty, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

TEOREMA 4.10

Si Y es una variable aleatoria exponencial con parámetro β, entonces

$$\mu = E(Y) = \beta \quad \text{y} \quad \sigma^2 = V(Y) = \beta^2.$$

Distribución de probabilidad t-Student

- Características:
  - Para validar si la diferencia entre las medias de dos muestras es estadísticamente significativa.
  - Tambien sirve para validar si dos variables están relacionadas

Distr. T-Student

$$p(t;n) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{2}\right)^{-\frac{n+1}{2}}$$