Distribuciones de probabilidad

Estadística Matemática

Dr. Jorge Samayoa
Ing. Preng Biba

Josué Benyamin Isaí Galeano Morales
Carnet: 18002955
II - FISICC

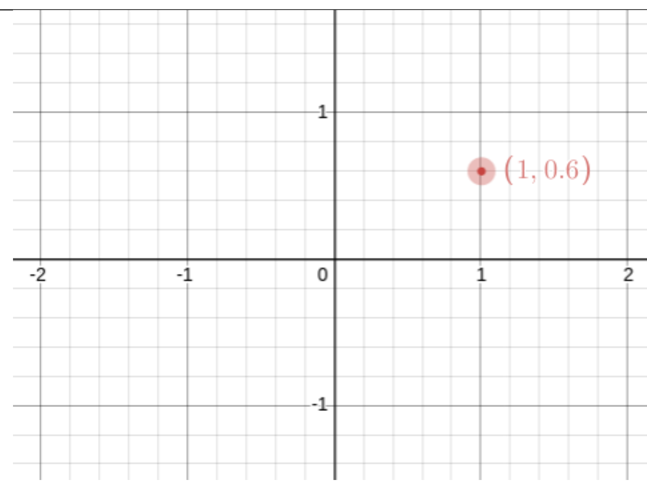
Guatemala, 10 de mayo de 2020

Índice

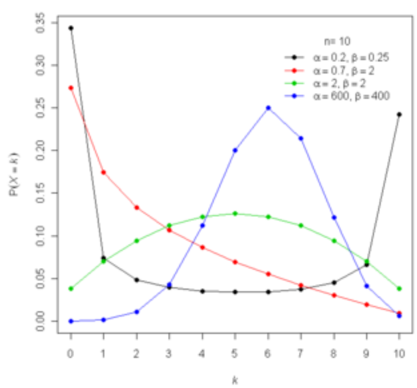
Distribuciones Discretas

Distribución de Bernoulli

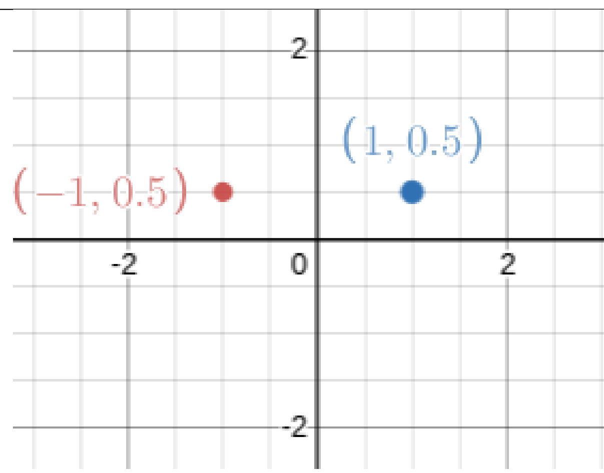
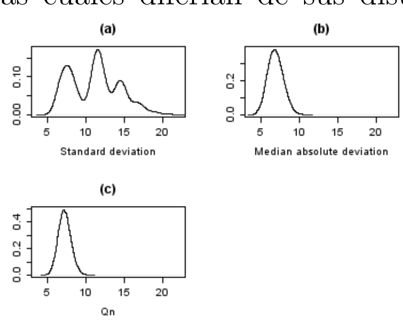
Es una distribución que sólo toma 2 valores, p cuando está evaluada en 1 y q cuando se evalúa en 0, la variable aleatoria en esta distribución cuenta la cantidad de éxitos, como está restringido su dominio en 0 y 1 significa que sólo puede decir cuanto es la probabilidad de que se falle o se tenga éxito, la distribución *Binomial* se basa en n pruebas independientes de la **distribución de Bernoulli**.

Resumen	
Variable Aleatoria	Cuenta la cantidad de éxitos
Func. de densidad de probabilidad	$f(x) = p^x(1-q)^{1-x}$, $x \in \{0, 1\}$
Func. de distribución acumulada	$f(0 < x) = 0, \quad f(0 \leq x < 1) = q, \quad f(x \geq 1) = 1$
Valor Esperado	$E(X) = p$
Varianza	$VAR(X) = pq$
Aplicaciones	Buscar la probabilidad cuando sólo se hace una prueba
Gráf. de la función de densidad	
Ejemplo	Suponemos que somos muy fans de un corredor de una competición ciclista en la cual solo compiten dos corredores. Queremos apostar a que ese corredor gana. $$p = 0.5, \quad q = 1 - p = 0.5$$ Entonces, si gana será un resultado "éxito" y si pierde será un resultado "no éxito".

Distribución Beta-Binomial

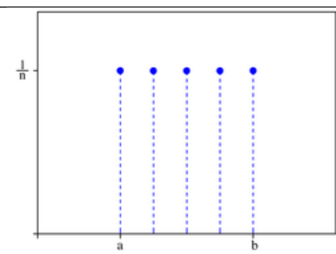
Resumen	
Variable Aleatoria	Número de éxitos, pero estos no tienen la misma probabilidad
Func. de densidad de probabilidad	$f(x) = \binom{n}{x}\frac{B(x+\alpha, n-x+\beta)}{B(\alpha,\beta)}$, donde $B(\alpha,\beta)$, es la función Beta
Func. de distribución acumulada	$f(0 \leq x < n) = \binom{n}{x}\frac{B(x+\alpha, n-x+\beta)}{B(\alpha,\beta)}\ {}_3F_2(a,b,x)$, esta es la función hipergeométrica generalizada, $f(0 < x) = 0,\ f(x \geq 1) = 1$
Valor Esperado	$E(X) = \dfrac{n\alpha}{\alpha+\beta}$
Varianza	$VAR(X) = \dfrac{n\alpha(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Aplicaciones	utilizada cuando se tiene condiciones similares que la *binomial* pero con probabilidad variable.
Gráf. de la función de densidad	
Ejemplo	Se quiere obtener la pdf de una variable aleatoria que cuenta el número de niños entre los primeros 12 jóvenes, de familias con 13 personas con una cantidad de 6115 familias para muestrear. La distribución binomial se queda muy por detrás de la beta-binomial ya no se dan más restricciones como para definir una probabilidad de éxito estática. datos: *boys* 0 1 2 3 4 5 6 ... *familias* 3 24 104 286 670 1033 1343 ...

Distribución de Rademacher

Resumen	
Variable Aleatoria	Nada en específico
Func. de densidad de probabilidad	$f(x) = 0.5 \; si \; x \in \{1, 2\}$, 0 en otro caso
Func. de distribución acumulada	$f(-1 < x) = 0, \; f(-1 \leq x < 1) = 0.5, \; f(x \geq 1) = 1$
Valor Esperado	$E(X) = 0$
Varianza	$VAR(X) = 1$
Aplicaciones	utilizada en bootstrapping, que es la práctica de estimar propiedades de un estimador
Gráf. de la función de densidad	
Ejemplo	Simon Newcomb la usó cuando obtuvo el conjunto de datos sobre la velocidad de la luz para la modelación de la desviación típica y mediana las cuales diferían de sus distribuciones muestrales. 

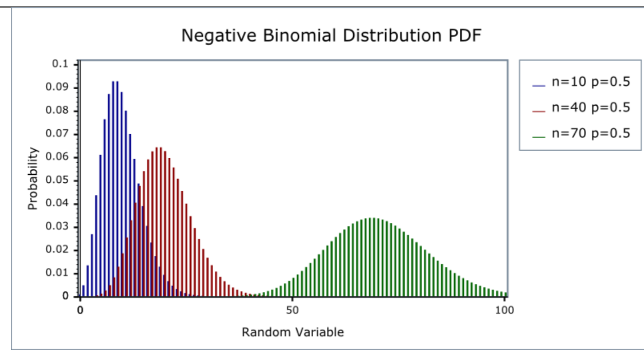
Distribución Uniforme discreta

Esta es equivalente a la ***Distribución normal continua*** sólo que con un número discreto de casos, por la forma en que está definida cada valor de la variable aleatoria está 1 entero de distancia del siguiente y el anterior.

Resumen	
Variable Aleatoria	modela eventos con la misma probabilidad
Func. de densidad de probabilidad	$f(x) = 1/n$
Func. de distribución acumulada	$F(x) = \dfrac{x-a+1}{n}$
Valor Esperado	$E(X) = \dfrac{a+b}{2}$
Varianza	$VAR(X) = \dfrac{(b-a+1)^2-1}{12}$
Aplicaciones	Para modelar la probabilidad de n resultados equiprobables (con la misma probabilidad de ocurrencia)
Gráf. de la función de densidad	
Ejemplo	se tira un dado balanceado, la probabilidad de que caiga un número en específico es de 1/6, cuando se tira una moneda balanceada, la probabilidad es de 1/2

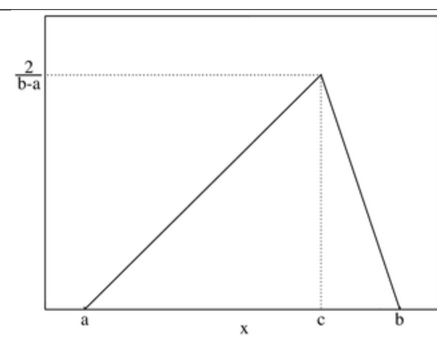
Distribución Binomial Negativa

Es una distribución compuesta de otras como las distribución de pascal (Geométrica) y la Binomial.

Resumen	
Variable Aleatoria	modela la cantidad de pruebas necesarias para obtener un número predefinido de r éxitos
Func. de densidad de probabilidad	$f(x) = \binom{x+r+1}{x}(1-p)^r p^x$
Func. de distribución acumulada	$F(x) = 1 - I_p(x+1, r)$, que es la función beta incompleta
Valor Esperado	$E(X) = \dfrac{pr}{1-p}$
Varianza	$VAR(X) = \dfrac{pr}{(1-p)^2}$
Aplicaciones	Cuando se necesita calcular cual es la probabilidad que ocurra r cantidad éxitos si se realizan dicha cantidad de pruebas.)
Gráf. de la función de densidad	
Ejemplo	Se requiere que Pat Collis venda golosinas para recaudar dinero para la excursión de 6to grado. Hay treinta casas en el vecindario, y no se supone que Pat regrese a casa hasta que se hayan vendido cinco barras de chocolate. Entonces el niño va de puerta en puerta y vende golosinas. En cada casa, hay una probabilidad de 0.6 de vender una barra de chocolate y una probabilidad de 0.4 de no vender nada. Como se puede observar se busca que se vendan 5 barras (éxitos).

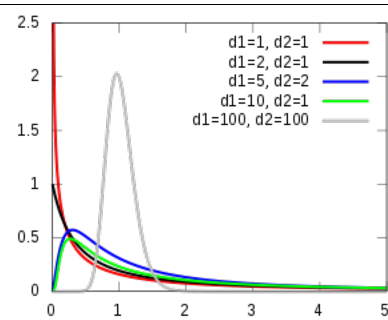
Distribuciones Continuas

Distribución Triangular

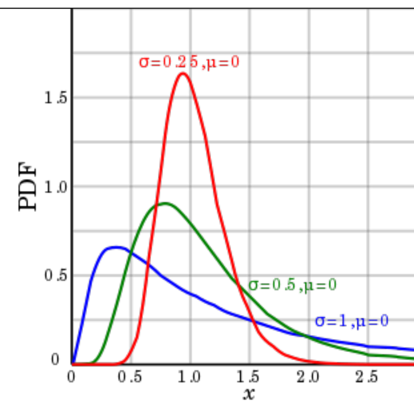
Resumen	
Variable Aleatoria	mide eventos donde se desconocen muchos datos.
Func. de densidad de probabilidad	$f(x) = 0, x < a \quad f(x) = \frac{2(x-a)}{(b-a)(c-a)}, a \leq x < c, \quad f(x) = \frac{2}{b-a}, x = c, \quad f(x) = \frac{2(b-x)}{(b-a)(b-c)}, c < x \leq b, \quad f(x) = 0, b < x$
Func. de distribución acumulada	$F(x) = 0, x \leq a \quad F(x) = \frac{(x-a)^2}{(b-a)(c-a)}, a < x \leq c, \quad F(x) = 1 - \frac{(b-x)^2}{(b-a)(b-c)}, c < x < b, \quad F(x) = 1, b \leq x$
Valor Esperado	$$\frac{a+b+c}{3}$$
Varianza	$$\frac{a^2+b^2+c^2-ab-ac-bc}{18}$$
Aplicaciones	habitualmente empleado para descripción subjetiva de poblaciones con cantidad limitada de datos
Gráf. de la función de densidad	
Ejemplo	se sabe que los efectos secundarios de ampliar una imagen con ancho 128px se distribuye triangular con $c = 35$ (la moda) y extremos $a = 0$ y $b = 128$

Nota: a, b son los extremos de la distribución y c es la moda.

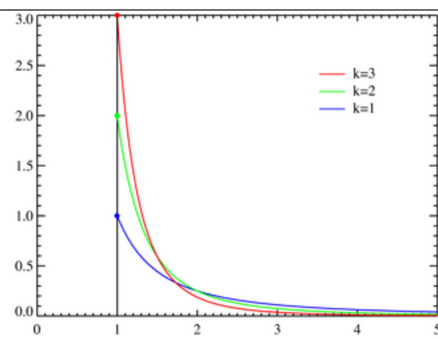
Distribución F

Resumen	
Variable Aleatoria	mide eventos donde se tiene un sesgo positivo (no es simétrica a su media)
Func. de densidad de probabilidad	$f(x) = \frac{\sqrt{\frac{(d_1 x)^{d_1} d_2{}^{d_2}}{(d_1 x + d_2)^{d_1+d_2}}}}{xB(\frac{d_1}{2} \frac{d_2}{2})}$, B es la función Beta.
Func. de distribución acumulada	$F(x) = I_{\frac{d_1 x}{d_1 x + d_2}}(\frac{d_1}{2}, \frac{d_2}{2})$
Valor Esperado	$\frac{d_2}{d_2-2} \quad d_2 > 2$
Varianza	$\frac{2d_2^2(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)} \quad d_2 > 4$
Aplicaciones	Pruebas de homocedasticidad
Gráf. de la función de densidad	
Ejemplo	Suponga que la mitad de los hogares en un país tienen ingresos inferiores a \$ 50,000 y la otra mitad tienen ingresos superiores a \$ 50,000; esto indica que el ingreso familiar promedio es de \$ 50,000. Entre los hogares con ingresos inferiores a \$ 50,000, el valor más pequeño posible es \$ 0. Entre los hogares con ingresos superiores a \$ 50,000, puede haber ingresos de varios millones de dólares por año. Este desequilibrio entre los ingresos por debajo de la mediana y por encima de la mediana hace que la media sea sustancialmente más alta que la mediana. Supongamos, por ejemplo, que el ingreso promedio en este caso es de \$ 120,000. Esto muestra que la distribución de los ingresos familiares está sesgada positivamente.

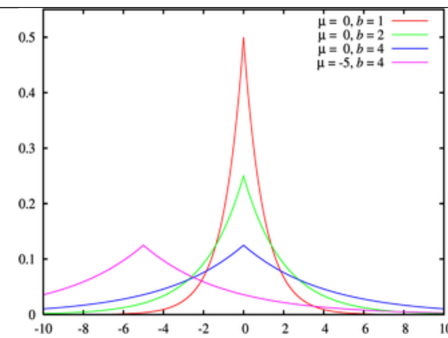
Distribución Lognormal

Resumen	
Variable Aleatoria	mide eventos con comportamiento simétrico en el logaritmo de su variable
Func. de densidad de probabilidad	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}$
Func. de distribución acumulada	$F(x) = \frac{1}{2} + \frac{1}{2} erf[\frac{lnx-\mu}{\sqrt{2}\sigma}]$, erf es la función de error.
Valor Esperado	$e^{\mu+\frac{\sigma^2}{2}}$
Varianza	$(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$
Aplicaciones	En la hidrología la utilizan para analizar variables aleatorias continuas como los valores máximos de la precipitación y la descarga de ríos, y para describir el comportamientos de las épocas de sequía.
Gráf. de la función de densidad	
Ejemplo	Cuando se toman datos de frecuencia estos la mayoría de veces se comportan con una distribución lognormal. Se estudió una empresa la cual cumplía tener parámetros $\mu = 4,\ \sigma = 2$, como esto se puede obtener la probabilidad de que dicha empresa esté usando cierta cantidad de potencia (db por hora), en cualquier hora peculiar.

Distribución de Pareto

Resumen	
Variable Aleatoria	mide eventos que se distribuyen de una forma parecida al logaritmo.
Func. de densidad de probabilidad	$f(x) = \dfrac{\alpha x_m^\alpha}{x^{\alpha+1}}\ x > x_m$
Func. de distribución acumulada	$F(x) = 1 - (\frac{x_m}{x})^\alpha,\ x \geq x_m$
Valor Esperado	$\dfrac{\alpha x_m}{\alpha - 1},\ \alpha > 1$
Varianza	$\dfrac{x_m^2 \alpha}{(\alpha-1)^2(\alpha-2)}$
Aplicaciones	Es utilizada al igual que la lognormal en la hidrología para analizar variables aleatorias continuas como los valores máximos de la precipitación y la descarga de ríos, y para describir el comportamientos de las épocas de sequía.
Gráf. de la función de densidad	
Ejemplo	algunos ejemplos de utilización de esta distribución se da en: El tamaño de los asentamientos humanos, distribución del tamaño del archivo del tráfico de Internet, Tasas de error de la unidad de disco duro.

Distribución de Laplace

Resumen	
Variable Aleatoria	mide eventos que tienen distribución en forma de dos exponenciales adyacentes y con caras contrarias.
Func. de densidad de probabilidad	$f(x) = \frac{1}{2b}e^{-\frac{\lvert x-\mu \rvert}{b}}$
Func. de distribución acumulada	$F(x) = 0.5e^{\frac{x-\mu}{b}}\ x \leq \mu, \quad F(x) = 1 - 0.5e^{\frac{x-\mu}{b}}\ x \geq \mu$
Valor Esperado	μ
Varianza	$2b^2$
Aplicaciones	Es utilizada al igual que la lognormal y de Pareto en la misma área con las mismas aplicaciones únicamente con la diferencia de como se ajustan.
Gráf. de la función de densidad	
Ejemplo	por ejemplo podemos utilizar la distribución de Laplace para el reconocimiento de voz y la compresión JPEG.

Referencias

[Dummies, 2020] Dummies (2020). F distribution - https://www.dummies.com/education/math/business-statistics/when-to-use-the-f-distribution/.

[Economipedia, 2020] Economipedia (2020). Distribución de bernoulli - https://economipedia.com/definiciones/ejemplo-de-distribucion-de-bernoulli.html.

[WikiMii, 2020] WikiMii (2020). Rademacher distribution - https://wikimili.com/en/Rademacher_distribution.

[Wikipedia, 2019a] Wikipedia (2019a). Distribución binomial negativa - https://es.wikipedia.org/w/index.php?title=Distribuci%C3%B3n_binomial_negativa&oldid=120619920.

[Wikipedia, 2019b] Wikipedia (2019b). Distribución de laplace - https://es.wikipedia.org/w/index.php?title=Distribuci%C3%B3n_de_Laplace&oldid=119627796.

[Wikipedia, 2019c] Wikipedia (2019c). Distribución triangular - https://en.wikipedia.org/wiki/Triangular_distribution#Business_simulations.

[Wikipedia, 2019] Wikipedia (2019). Rademacher distribution - https://en.wikipedia.org/w/index.php?title=Rademacher_distribution&oldid=920903351.

[Wikipedia, 2020] Wikipedia (2020). Beta-binomial distribution - https://en.wikipedia.org/w/index.php?title=Beta-binomial_distribution&oldid=953226575.

[Wikipedia, 2020] Wikipedia (2020). Bootstrapping - https://es.wikipedia.org/wiki/Bootstrapping_(estad%C3%ADstica).

[Wikipedia, 2020] Wikipedia (2020). Discrete uniform distribution - https://en.wikipedia.org/w/index.php?title=Discrete_uniform_distribution&oldid=945319081.

[Wikipedia, 2020a] Wikipedia (2020a). Distribución de bernoulli - https://es.wikipedia.org/w/index.php?title=Distribuci%C3%B3n_de_Bernoulli&oldid=125659614.

[Wikipedia, 2020b] Wikipedia (2020b). Distribución lognormal - https://en.wikipedia.org/wiki/Log-normal_distribution.

[Wikipedia, 2020c] Wikipedia (2020c). Distribución pareto - https://en.wikipedia.org/wiki/Pareto_distribution.

[Wikipedia, 2020d] Wikipedia (2020d). Tramado - https://es.wikipedia.org/wiki/Tramado.