

Variables Aleatorias

Estadística Computacional

Juan Zamora Osorio
juan.zamora@pucv.cl

Instituto de Estadística
Pontificia Universidad Católica de Valparaíso

18 de octubre de 2023



PONTIFICIA
UNIVERSIDAD
CATÓLICA DE
VALPARAÍSO

Variables Aleatorias – cantidad de cartas negras

Sacando una carta – posibilidades

♣	1
♦	0
♥	0
♠	1

Sacando una carta – probabilidades

Cantidad negras	0	1
Probabilidad	$\frac{2}{4} = \frac{1}{2}$	$\frac{2}{4} = \frac{1}{2}$

Variables Aleatorias – cantidad de cartas negras

Sacando dos cartas – posibilidades

	♣	♦	♥	♠
♣	2	1	1	2
♦	1	0	0	1
♥	1	0	0	1
♠	2	1	1	2

Sacando dos cartas – probabilidades

Cantidad negras	0	1	2
Probabilidad	$\frac{4}{16} = \frac{1}{4}$	$\frac{8}{16} = \frac{1}{2}$	$\frac{4}{16} = \frac{1}{4}$

Variables Aleatorias – cantidad de cartas negras

Sacando tres cartas – posibilidades

	♣	♦	♥	♠	♦	♣	♦	♥	♠
♣	3	2	2	3	♣	2	1	1	2
♦	2	1	1	2	♦	1	0	0	1
♥	2	1	1	2	♥	1	0	0	1
♠	3	2	2	3	♠	2	1	1	2
♥	2	1	1	2	♠	3	2	2	3
♣	2	1	1	2	♣	2	1	1	2
♦	1	0	0	1	♦	2	1	1	2
♥	1	0	0	1	♥	2	1	1	2
♠	2	1	1	2	♠	3	2	2	3

Sacando tres cartas – probabilidades

Cantidad negras	0	1	2	3
Probabilidad	$\frac{8}{64} = \frac{1}{8}$	$\frac{24}{64} = \frac{3}{8}$	$\frac{24}{64} = \frac{3}{8}$	$\frac{8}{64} = \frac{1}{8}$

Variables Aleatorias – definición

Formalmente

- ▶ Función *medible* entre un espacio muestral Ω y otro Ω' :
$$X : \Omega \rightarrow \Omega'.$$
- ▶ La preimagen de todo conjunto en el espacio de eventos asociado a Ω' es un conjunto en el espacio de eventos de Ω .

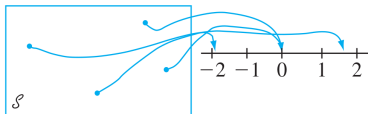
Ejemplo

- ▶ X cantidad de cartas negras al sacar 3 cartas independientes.
- ▶ $\Omega' = \{0, 1, 2, 3\}$.
- ▶ $\Omega =$
 $\{(\clubsuit, \clubsuit, \clubsuit), (\clubsuit, \clubsuit, \diamondsuit), (\clubsuit, \clubsuit, \heartsuit), (\clubsuit, \clubsuit, \spadesuit), (\clubsuit, \diamondsuit, \clubsuit), \dots, \}.$
- ▶ $X((\clubsuit, \diamondsuit, \clubsuit)) = 2.$

En este curso

$$X : \Omega \rightarrow E$$

- ▶ Si E es finito o numerable: variable aleatoria discreta.
- ▶ Si $E \subseteq \mathbb{R}$ no numerable: variable aleatoria continua.



Ejemplo – variable aleatoria discreta

- ▶ X cantidad de cartas negras al sacar 3 cartas independientes.

Cantidad negras	Probabilidad
0	0,125
1	0,375
2	0,375
3	0,125

Notación

$$X : \Omega \rightarrow E$$

- ▶ Sea $S \subseteq E$, $P_X(S) = P(X \in S) = P(X^{-1}(S)) = P(\{\omega \in \Omega : X(\omega) \in S\})$.
- ▶ P_X :
 - ▶ Distribución de probabilidad.
 - ▶ Ley de probabilidad.

Ejemplo – variable aleatoria discreta

- ▶ $P_X(\{0, 1\}) =$
 $P(\{\diamond\diamond\diamond, \diamond\heartsuit\clubsuit, \spadesuit\diamond\heartsuit, \spadesuit\heartsuit\heartsuit, \diamond\clubsuit\heartsuit, \diamond\diamond\heartsuit, \dots\})$.

X	P_X
0	0,125
1	0,375
2	0,375
3	0,125

Variables aleatorias discretas

Función de masa de probabilidad (*pmf*)

$$f(x) = P(X = x) = P_X(\{x\}).$$

Recordar de probabilidades (Ω finito)

$$f(x) = P(X = x) = \frac{n_x}{n} = \frac{|X^{-1}(\{x\})|}{|\Omega|}.$$

- ▶ n_x cantidad de resultados ω que cumplen $X(\omega) = x$.
- ▶ n cantidad de resultados en Ω .

Ejemplo: moneda lanzada 3 veces

Ejemplo 1

- ▶ X cantidad de caras.
- ▶ Posibles valores: $\{0, 1, 2, 3\}$.

x	$X^{-1}(\{x\})$	$P(X = x)$
0	$\{sss\}$	$\frac{1}{8} = 0,125$
1	$\{css, scs, ssc\}$	$\frac{3}{8} = 0,375$
2	$\{ccs, csc, scc\}$	$\frac{3}{8} = 0,375$
3	$\{ccc\}$	$\frac{1}{8} = 0,125$

Ejemplo: moneda lanzada 3 veces

Ejemplo 2

- X string de bits 0/1 según cara/sello obtenido.
- Posibles valores: $\{000, 001, 010, 011, 100, 101, 110, 111\}$.

x	$X^{-1}(\{x\})$	$P(X = x)$
000	$\{ccc\}$	$\frac{1}{8} = 0,125$
001	$\{ccs\}$	$\frac{1}{8} = 0,125$
010	$\{csc\}$	$\frac{1}{8} = 0,125$
011	$\{css\}$	$\frac{1}{8} = 0,125$
100	$\{scc\}$	$\frac{1}{8} = 0,125$
101	$\{scs\}$	$\frac{1}{8} = 0,125$
110	$\{ssc\}$	$\frac{1}{8} = 0,125$
111	$\{sss\}$	$\frac{1}{8} = 0,125$

Ejemplo: moneda lanzada 3 veces

Ejemplo 3

- ▶ X cantidad de caras par o impar.
- ▶ Posibles valores: $\{P, I\} \equiv \{0, 1\}$.

x	$X^{-1}(\{x\})$	$P(X = x)$
$P \equiv 0$	$\{ccs, csc, scc, sss\}$	$\frac{1}{2} = 0,5$
$I \equiv 1$	$\{ccc, css, scs, ssc\}$	$\frac{1}{2} = 0,5$

Función de distribución acumulada (*cdf*)

Definición

$F : \mathbb{R} \rightarrow [0, 1]$ tal que $F(x) = P(X \leq x)$.

Variables aleatorias discretas

$$F(x) = P(X \leq x) = \sum_{y \leq x} P(X = y) = \sum_{y \leq x} f(y).$$

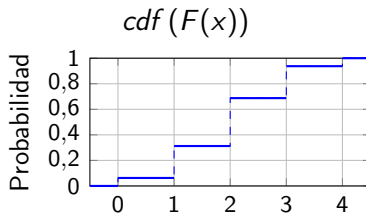
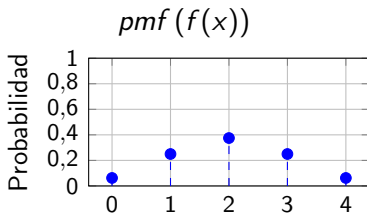
Propiedades

- ▶ F es no decreciente ($x_1 \leq x_2 \Rightarrow F(x_1) \leq F(x_2)$).
- ▶ $\lim_{x \rightarrow -\infty} F(x) = 0$.
- ▶ $\lim_{x \rightarrow +\infty} F(x) = 1$.
- ▶ $\forall x, 0 \leq F(x) \leq 1$.
- ▶ Está definida para todo \mathbb{R} .

Función de distribución acumulada (cdf)

Ejemplo: Cantidad de caras al lanzar 4 monedas

x	$f(x) = P(X = x)$	$F(x) = \sum_{y < x} f(y)$
0	$\frac{1}{16} = 0,0625$	$\frac{1}{16} = 0,0625$
1	$\frac{4}{16} = 0,25$	$\frac{5}{16} = 0,3125$
2	$\frac{6}{16} = 0,375$	$\frac{11}{16} = 0,6875$
3	$\frac{4}{16} = 0,25$	$\frac{15}{16} = 0,9375$
4	$\frac{1}{16} = 0,0625$	$\frac{16}{16} = 1$



Esperanza – valor esperado – media

Definición

$$\mathbb{E}_{X \sim f}[X] = \mu_X = \sum_{x \in E} xf(x).$$

Ejemplo: Cantidad de caras al lanzar 4 monedas

$$\mathbb{E}[X] = \frac{1}{16}0 + \frac{4}{16}1 + \frac{6}{16}2 + \frac{4}{16}3 + \frac{1}{16}4 = 2.$$

Esperanza – valor esperado – media

Definición

- ▶ Sea $g : E \rightarrow \mathbb{R}$ una función:

$$\mathbb{E}_{X \sim f}[g(X)] = \sum_{x \in E} g(x)f(x).$$

Propiedades: linealidad

- ▶ Sean $a, b \in \mathbb{R}$:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b.$$

- ▶ Sean $g_1, g_2 : E \rightarrow \mathbb{R}$ funciones:

$$\mathbb{E}[g_1(X) + g_2(X)] = \mathbb{E}[g_1(X)] + \mathbb{E}[g_2(X)].$$

Varianza

Definición

$$\mathbb{V}_{X \sim f}[X] = \sigma_X^2 = \mathbb{E}_{X \sim f}[(X - \mu_X)^2] = \sum_{x \in E} (x - \mu_X)^2 f(x).$$

Desviación estándar

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{\mathbb{V}_{X \sim f}[X]}.$$

Propiedad

$$\mathbb{V}_{X \sim f}[X] = \sigma_X^2 = \mathbb{E}_{X \sim f}[X^2] - \mathbb{E}_{X \sim f}[X]^2 = \left[\sum_{x \in E} x^2 f(x) \right] - \mu_X^2.$$

Varianza

Definición

- Sea $g : E \rightarrow \mathbb{R}$ una función:

$$\mathbb{V}_{X \sim f}[g(X)] = \sum_{x \in E} (g(x) - \mathbb{E}_{X \sim f}[g(X)])^2 f(x).$$

Propiedad: no linealidad

- Sean $a, b \in \mathbb{R}$:

$$\mathbb{V}[aX + b] = a^2 \mathbb{V}[X].$$

Distribución $X \sim \text{Bernoulli}(p)$

Definición

$$f(x) = \begin{cases} p & \text{si } x = 1, \\ 1 - p & \text{si } x = 0. \end{cases}$$

Ejemplos

- ▶ Lanzar una moneda.
- ▶ Ser fiscalizado en carretera.

Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Bernoulli}(p)}[X] = \mu_X = p.$
- ▶ $\mathbb{V}_{X \sim \text{Bernoulli}(p)}[X] = \sigma_X^2 = p(1 - p).$

Ejemplo: Reunión de personas sin COVID-19

Supuestos

- ▶ Reunión de personas, van llegando de a una.
- ▶ Cada persona que llega tiene probabilidad p de tener COVID-19.
- ▶ Probabilidad p es independiente entre personas.
- ▶ $X - 1$ cantidad de personas sanas cuando llega primera con COVID-19 (Igual cantidad total de personas menos 1)

x	Evento	$f(x) = P(X = x)$	$F(x) = \sum_{y \leq x} f(y)$
1	$\{c \dots\}$	p	p
2	$\{nc \dots\}$	$(1 - p)p$	$(2 - p)p$
3	$\{nnc \dots\}$	$(1 - p)^2 p$	$(3 - 3p + p^2)p$
4	$\{nnnc \dots\}$	$(1 - p)^3 p$	$(4 - 6p + 4p^2 - p^3)p$
\vdots	\vdots	\vdots	\vdots

Ejemplo: Reunión de personas sin COVID-19

x	Evento	$f(x) = P(X = x)$	$F(x) = \sum_{y \leq x} f(y)$
1	$\{c \dots\}$	p	p
2	$\{nc \dots\}$	$(1 - p)p$	$(2 - p)p$
3	$\{nnc \dots\}$	$(1 - p)^2 p$	$(3 - 3p + p^2)p$
4	$\{nnnc \dots\}$	$(1 - p)^3 p$	$(4 - 6p + 4p^2 - p^3)p$
\vdots	\vdots	\vdots	\vdots

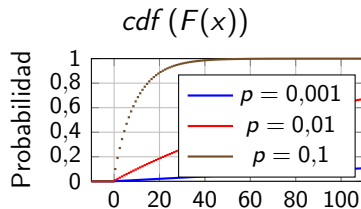
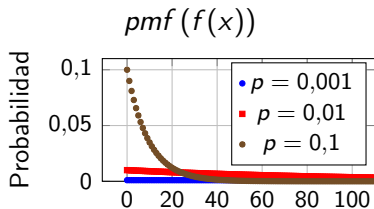
En general

- ▶ $f(x) = P(X = x) = (1 - p)^{x-1} p.$
- ▶ $F(x) = P(X \leq x) = 1 - (1 - p)^x.$

Ejemplo: Reunión de personas sin COVID-19

En general

- ▶ $f(x) = P(X = x) = (1 - p)^{x-1}p$.
- ▶ $F(x) = P(X \leq x) = 1 - (1 - p)^x$.



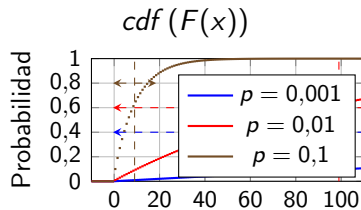
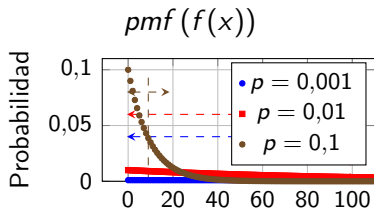
¿Qué pasó?

- ▶ X ahora está expresada como una función pmf en vez de una tabla.

Distribución $X \sim \text{Geométrica}(p)$

Definición

$$f(x) = P(X = x) = (1 - p)^{x-1}p, \text{ con } x \in \{0, 1, 2, \dots\}.$$



Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Geométrica}(p)}[X] = \mu_X = \frac{1}{p}.$
- ▶ $\mathbb{V}_{X \sim \text{Geométrica}(p)}[X] = \sigma_X^2 = \frac{1-p}{p^2}.$

Ejemplo: Reunión de personas sin COVID-19

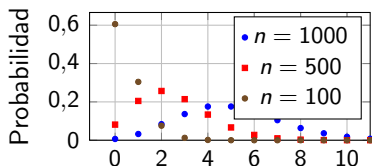
Supuestos

- ▶ Cada persona que llega tiene prob. p de tener COVID-19.
- ▶ Probabilidad p es independiente entre personas.
- ▶ X personas con COVID-19 en reunión de n personas.

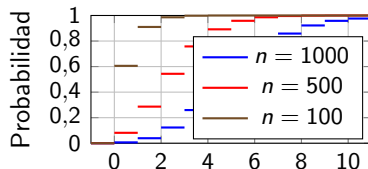
X como función *pmf*

$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

pmf ($f(x)$) con $p = 0,005$



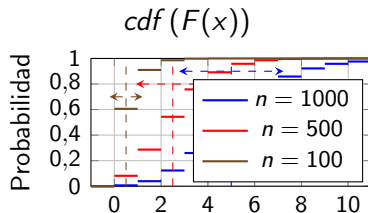
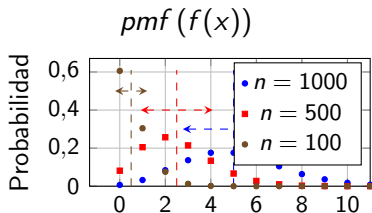
cdf ($F(x)$) con $p = 0,005$



Distribución $X \sim \text{Binomial}(n, p) = B(n, p)$

Definición

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ con } x \in \{0, 1, \dots, n\}.$$



Propiedades

- ▶ $\mathbb{E}_{X \sim B(n,p)}[X] = \mu_X = np.$
- ▶ $\mathbb{V}_{X \sim B(n,p)}[X] = \sigma_X^2 = np(1-p).$

Distribución $X \sim \text{Poisson}(\lambda) = \text{Pois}(\lambda)$

Definición

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \text{ con } x \in \{0, 1, \dots\}.$$

Ejemplos

- ▶ Cantidad de autos que pasan frente a la casa cada una hora.
- ▶ Cantidad de fotones que llegan a un pixel durante fotografía.

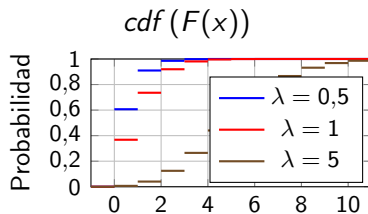
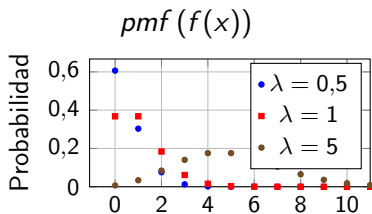
Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Pois}(\lambda)}[X] = \mu_X = \lambda.$
- ▶ $\mathbb{V}_{X \sim \text{Pois}(\lambda)}[X] = \sigma_X^2 = \lambda.$

¿De dónde viene?

- ▶ Si $X \sim B(n, p)$, el valor esperado es np .
- ▶ Fijando el valor esperado, se deja crecer $n \rightarrow +\infty$.

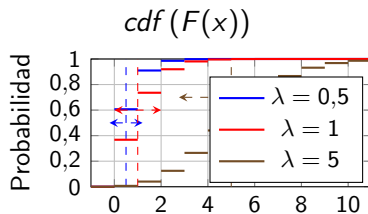
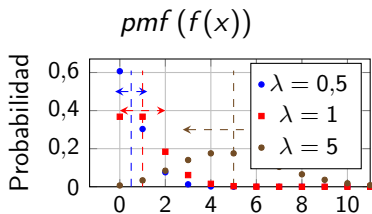
Distribución $X \sim \text{Poisson}(\lambda) = \text{Pois}(\lambda)$



Propiedades

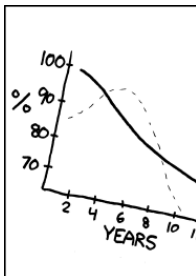
- ▶ $\mathbb{E}_{X \sim \text{Pois}(\lambda)}[X] = \mu_X = \lambda.$
- ▶ $\mathbb{V}_{X \sim \text{Pois}(\lambda)}[X] = \sigma_X^2 = \lambda.$

Distribución $X \sim \text{Poisson}(\lambda) = \text{Pois}(\lambda)$

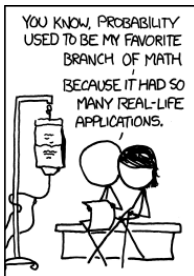


Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Pois}(\lambda)}[X] = \mu_X = \lambda.$
- ▶ $\mathbb{V}_{X \sim \text{Pois}(\lambda)}[X] = \sigma_X^2 = \lambda.$



5 YEARS	81%
10 YEARS	77%



Variables aleatorias continuas

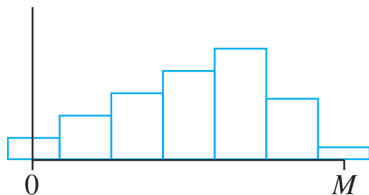
Ejemplo

- ¿Cómo se distribuyen los tiempos de espera en la fila para almorzar?

Solución 1: discretizar en clases intervalos de ancho $a = 1$

- $\text{Clases}_{a=1} = \{[0, 1], (1, 2], (2, 3], \dots\}$.
- Se cumple que:

$$P(X \in [0, 1]) + \sum_{i=1}^{+\infty} P(X \in (i, i + 1]) = 1.$$



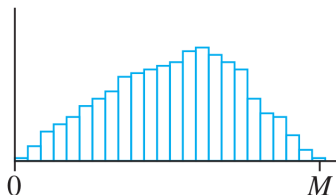
Variables aleatorias continuas

Solución $\frac{1}{2}$: discretizar en clases intervalos de ancho $a = \frac{1}{2}$

► $\text{Clases}_{a=\frac{1}{2}} = \left\{ \left[0, \frac{1}{2}\right], \left(\frac{1}{2}, 1\right], \left(1, \frac{3}{2}\right], \dots \right\}$.

► Se cumple que:

$$P\left(X \in \left[0, \frac{1}{2}\right]\right) + \sum_{i=1}^{+\infty} P\left(X \in \left(\frac{i}{2}, \frac{i+1}{2}\right]\right) = 1.$$



Variables aleatorias continuas

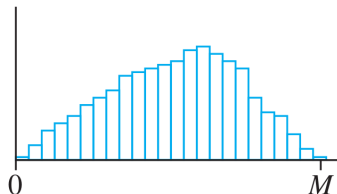
Solución a

- ▶ Clases_a = $\{[0, a], (a, 2a], (2a, 3a], \dots\}$.
- ▶ Se cumple que:

$$P(X \in [0, a]) + \sum_{i=1}^{+\infty} P(X \in (ia, (i+1)a]) = 1.$$

- ▶ Recordando que $P(X \leq x) = F(x)$, se puede reescribir como:

$$F(a) + \sum_{i=1}^{+\infty} F((i+1)a) - F(ia) = 1.$$



Variables aleatorias continuas

Solución a

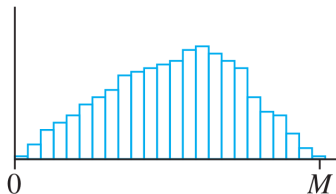
- Se tiene que:

$$F(a) + \sum_{i=1}^{+\infty} F((i+1)a) - F(ia) = F(a) + \sum_{i=1}^{+\infty} \text{Área}(i) = 1,$$

con $\text{Área}(i) = h_i a = P(X \in (ia, (i+1)a]) = F((i+1)a) - F(ia)$.

- Notar que la altura cumple:

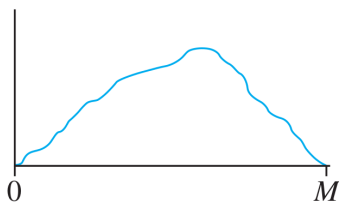
$$h_i = \frac{F(ia + a) - F(ia)}{a}.$$



Variables aleatorias continuas

Solución $a \rightarrow 0$

$$h_x = \lim_{a \rightarrow 0} \frac{F(x+a) - F(x)}{a}, \text{ y } \int_{x \in E} h_x dx = 1.$$



Función de densidad de probabilidad (*pdf*)

- ▶ $p(x) \equiv h_x$.
- ▶ Debe cumplir:
 - ▶ $\forall x \in E, p(x) \geq 0$. $\int_{x \in E} p(x) dx = 1$.

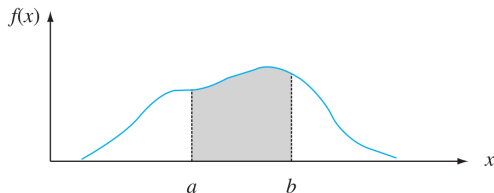
Variables aleatorias continuas – definiciones

Función de densidad de probabilidad (*pdf*)

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

Función de distribución acumulada (*cdf*)

$$F(x) = P(X \leq x) = \int_{-\infty}^x p(t) dt.$$



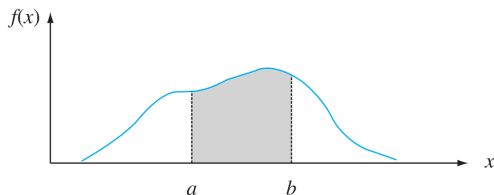
Variables aleatorias continuas – definiciones

Valor esperado / esperanza

$$\mathbb{E}_{X \sim p}[X] = \mu_X = \int_{x \in E} xp(x)dx.$$

Varianza

$$\mathbb{V}_{X \sim p}[X] = \sigma_X^2 = \int_{x \in E} (x - \mathbb{E}[X])^2 p(x) dx.$$

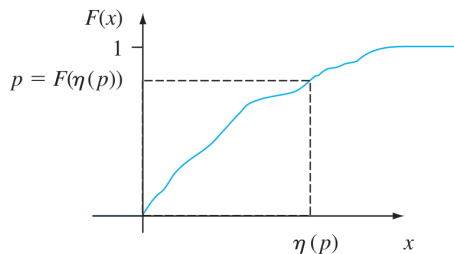
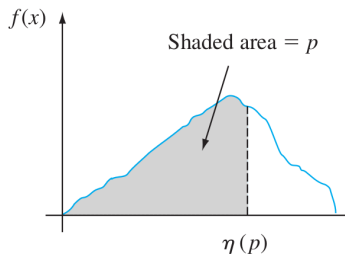


Variables aleatorias continuas – definiciones

Cuantil / percentil

► q_u tal que

$$P(X \leq q_u) = F(q_u) = \int_{-\infty}^{q_u} p(x) dx = u.$$



► En esta figura, $f(x) = p(x)$, $p = u$, $\eta(p) = q_u$.

Distribución $X \sim \text{Uniforme}(a, b) = U(a, b)$

Ángulos

- ▶ Ángulo luego de girar ruleta.
- ▶ Ángulo desde el centro al lanzar un dardo.

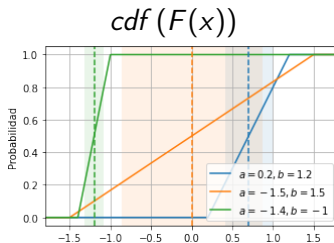
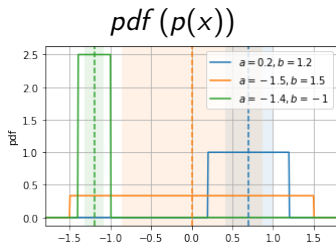
Otros

- ▶ Valor del parámetro p en una distribución Bernoulli(p).

Distribución $X \sim \text{Uniforme}(a, b) = U(a, b)$

Definición

$$p(x) = \frac{1}{b-a}, \text{ con } x \in [a, b].$$



Distribución $X \sim \text{Uniforme}(a, b) = U(a, b)$

Propiedades

- ▶ $\mathbb{E}_{X \sim U(a,b)}[X] = \mu_X = \frac{a+b}{2}.$
- ▶ $\mathbb{V}_{X \sim U(a,b)}[X] = \sigma_X^2 = \frac{(b-a)^2}{12}.$

Ejemplo – tiempo entre vacunados

Supuesto

- ▶ Número de personas que llegan a vacunatorio sigue distribución Poisson con media 12 personas por hora.
- ▶ Luego de llegar al vacunatorio, ¿cuál es la probabilidad de que la próxima persona llegue
 - ▶ en más de 1 hora?
 - ▶ en más de 15 minutos?
 - ▶ en más de 5 minutos?
 - ▶ en menos de 5 minutos?

Recordar

$$X \sim \text{Pois}(\lambda) \Rightarrow P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

Ejemplo – tiempo entre vacunados

Supuesto

- ▶ Número de personas que llegan a vacunatorio sigue distribución Poisson con media 12 personas por hora.
- ▶ Luego de llegar al vacunatorio, ¿cuál es la probabilidad de que la próxima persona llegue...

en más de 1 hora?

$$X \sim \text{Pois}(12) \Rightarrow P(X = x) = e^{-12} \frac{12^x}{x!} \Rightarrow P(X = 0) = e^{-12}.$$

en más de 15 minutos?

- ▶ 12 personas por hora es equivalente a 3 persona cada 15 minutos.

$$X \sim \text{Pois}(3) \Rightarrow P(X = x) = e^{-3} \frac{3^x}{x!} \Rightarrow P(X = 0) = e^{-3}.$$

Ejemplo – tiempo entre vacunados

Supuesto

- ▶ Número de personas que llegan a vacunatorio sigue distribución Poisson con media 12 personas por hora.
- ▶ Luego de llegar al vacunatorio, ¿cuál es la probabilidad de que la próxima persona llegue...

en más de 5 minutos?

- ▶ 12 personas por hora es equivalente a 1 persona cada 5 minutos.

$$X \sim \text{Pois}(1) \Rightarrow P(X = x) = e^{-1} \frac{1}{x!} \Rightarrow P(X = 0) = e^{-1}.$$

en menos de 5 minutos?

- ▶ 12 personas por hora es equivalente a 1 persona cada 5 minutos.

$$P(X = x) = e^{-1} \frac{1}{x!} \Rightarrow P(X > 0) = 1 - P(X = 0) = 1 - e^{-1}.$$

Ejemplo – tiempo entre vacunados

En general

- Probabilidad de que la próxima persona llegue en más de t horas.

$$X \sim \text{Pois}(\lambda t) \Rightarrow P(X = x) = e^{-\lambda t} \frac{(\lambda t)^x}{x!} \Rightarrow P(X = 0) = e^{-\lambda t}.$$

- Probabilidad de que la próxima persona llegue en menos de t horas.

$$X \sim \text{Pois}(\lambda t) \Rightarrow P(X = x) = e^{-\lambda t} \frac{(\lambda t)^x}{x!} \Rightarrow P(X > 0) = 1 - e^{-\lambda t}.$$

¿Qué pasó?

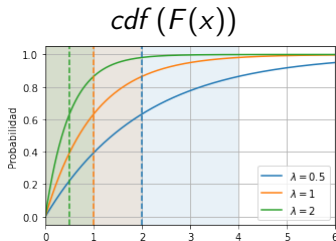
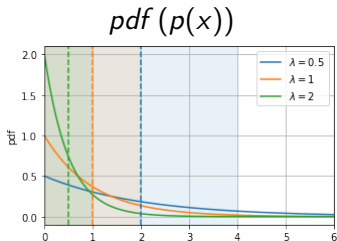
- ▶ Estas probabilidades son funciones continuas de t .
- ▶ $e^{-\lambda 0} = 1$ y $\lim_{t \rightarrow +\infty} e^{-\lambda t} = 0$.
- ▶ Supongamos que existe una distribución continua de probabilidad del tiempo T de llegada de siguiente persona:

$$P(T \leq t) = F(t) = \int_{u=0}^t p(u) du = 1 - e^{-\lambda t}.$$

Distribución Exponencial(λ)

Solución

$$p(x) = \lambda e^{-\lambda x} \approx \frac{1}{\beta} e^{-\frac{x}{\beta}}$$



Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Exponencial}(\lambda)}[X] = \mu_X = \frac{1}{\lambda}$.
- ▶ $\mathbb{V}_{X \sim \text{Exponencial}(\lambda)}[X] = \sigma_X^2 = \frac{1}{\lambda^2}$.

Distribución Normal $(0, \sigma^2) = \mathcal{N}(0, \sigma^2)$

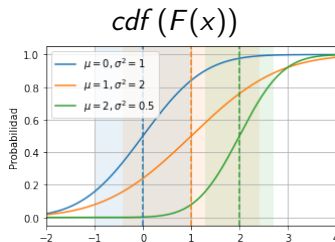
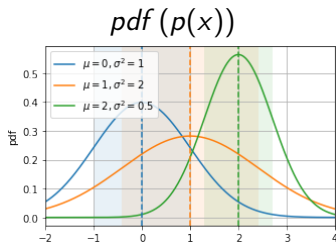
Propiedades

- ▶ $\mathbb{E}_{X \sim \mathcal{N}(0, \sigma^2)}[X] = \mu_X = 0.$
- ▶ $\mathbb{V}_{X \sim \mathcal{N}(0, \sigma^2)}[X] = \sigma_X^2 = \sigma^2.$

Distribución Normal(μ, σ^2) = $\mathcal{N}(\mu, \sigma^2)$

Definición

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

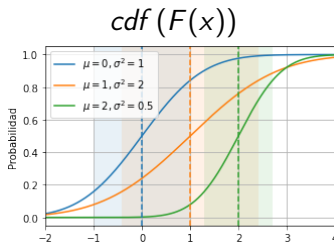
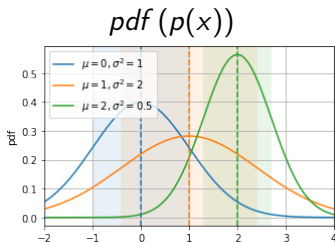


Distribución Normal(μ, σ^2) = $\mathcal{N}(\mu, \sigma^2)$

Definición

$$p(x) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}}_{\text{constante de normalización}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

constante de normalización



Función de distribución acumulada (*cdf*)

- No se puede escribir la *cdf* directamente.

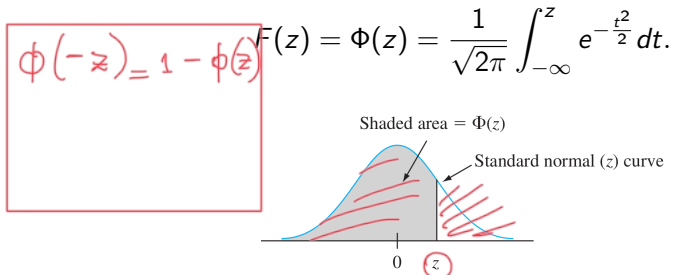
Distribución Normal(0, 1) = $\mathcal{N}(0, 1)$ estándar

Estandarizando

$$Z = \frac{X - \mu}{\sigma} \Rightarrow p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

► $\mu_Z = 0$ y $\sigma_Z^2 = 1$.

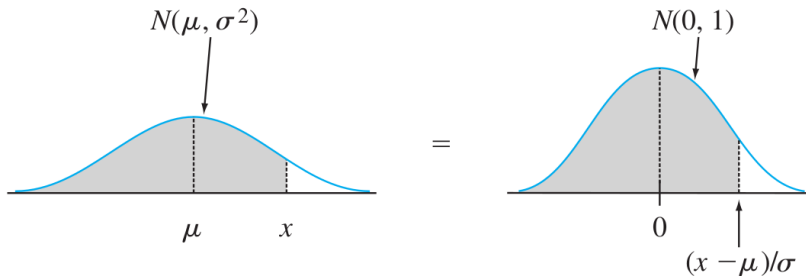
Función de distribución acumulada (cdf) $\mathcal{N}(0, 1)$ – definición



Distribución Normal(μ, σ^2) = $\mathcal{N}(\mu, \sigma^2)$

Volviendo a X : función de distribución acumulada (*cdf*)
 $\mathcal{N}(\mu, \sigma^2)$

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

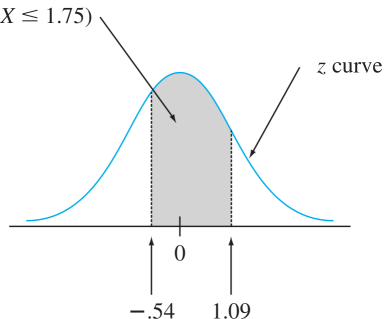
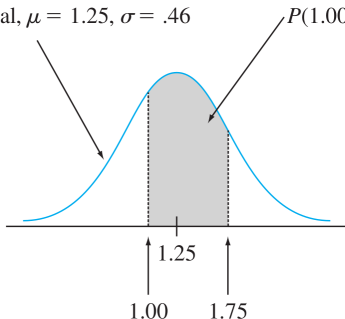


Distribución Normal(μ, σ^2) = $\mathcal{N}(\mu, \sigma^2)$

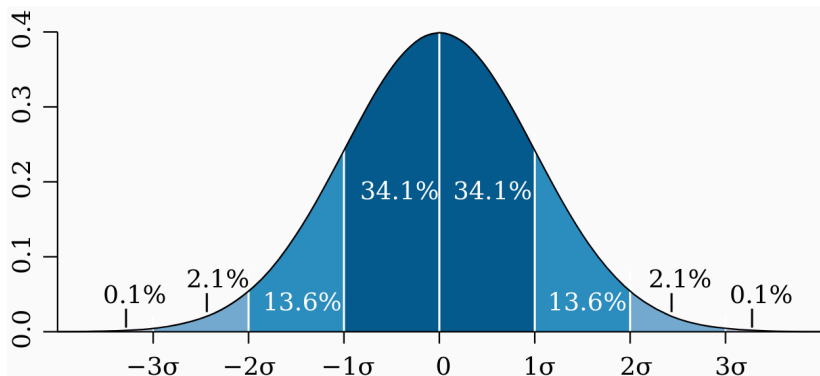
Probabilidad en intervalo $[a, b]$

$$P(a \leq X \leq b) = F(b) - F(a) = \Phi\left(\frac{\overset{b!}{b - \mu}}{\sigma}\right) - \Phi\left(\frac{\overset{a!}{a - \mu}}{\sigma}\right).$$

Normal, $\mu = 1.25, \sigma = .46$



Distribución Normal(μ, σ^2) = $\mathcal{N}(\mu, \sigma^2)$



Distribución normal

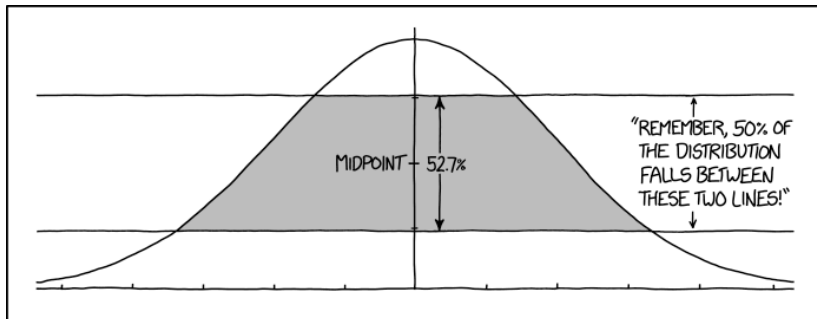
Ejemplo: altura personas en Chile

- ▶ Se ha determinado que la altura media de las mujeres es de 159,4 cm.
- ▶ Se ha determinado que la altura media de los hombres es de 171,8 cm.
- ▶ Suponiendo que la altura se distribuye como una normal:
 - ▶ ¿Es un hombre que mide 165 cm. bajo?
 - ▶ ¿Es una mujer que mide 165 cm. alta?

Respuesta

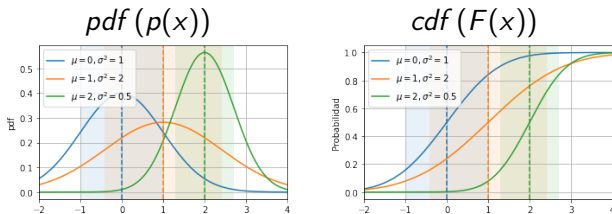
- ▶ ¡Depende de la desviación estándar / varianza!

Distribución normal



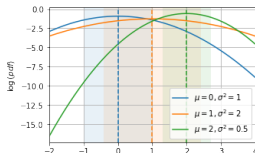
HOW TO ANNOY A STATISTICIAN

Distribución Normal(μ, σ^2) = $\mathcal{N}(\mu, \sigma^2)$



Tal vez ayuda para recordar

$$\ln(p(x)) = -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x - \mu)^2}{2\sigma^2}.$$



¿Por qué usar distribución $\mathcal{N}(\mu, \sigma^2)$?

- ▶ Termodinámica: distribución tiene la mayor entropía, sujeto a μ y σ^2 .
- ▶ Igualmente, distribución con menos información además de μ y σ^2 .
- ▶ La suma de distribuciones converge a una distribución normal.
- ▶ Logaritmo de producto de distribuciones converge a una normal.



Sobre constante de normalización

Recordar

$$p(x) = \frac{1}{\underbrace{\sqrt{2\pi\sigma^2}}_{\text{constante de normalización}}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Sobre constante de normalización

En general

- ▶ Muchas veces se define la “forma” de una distribución.
- ▶ Ejemplos:
 - ▶ $p(x) \propto 1$.
 - ▶ $p(x) \propto e^{-x^2}$.
 - ▶ $p(x) \propto e^{-x}$, con $x \geq 0$.
 - ▶ $p(x) \propto e^{-|x|}$.
- ▶ Luego se calcula la constante de normalización para que sea una distribución de probabilidad:

$$\int_{-\infty}^{\infty} p(x) dx = 1.$$

Sobre constante de normalización

Ejemplo

$$p(x) \propto e^{-\lambda x}, \text{ con } x \geq 0, \lambda > 0.$$

Desarrollo

$$\int_0^{\infty} p(x) dx = \int_0^{\infty} ce^{-\lambda x} dx = -c \frac{e^{-\lambda x}}{\lambda} \Big|_{x=0}^{\infty} = \frac{c}{\lambda} = 1 \Rightarrow c = \lambda.$$

Entonces

$$p(x) = \lambda e^{-\lambda x}, \text{ con } x \geq 0, \lambda > 0.$$

Sobre constante de normalización

Ejemplo

$$p(x) \propto e^{-\frac{|x-\mu|}{b}}, \text{ con } b > 0.$$

Desarrollo

$$\begin{aligned} \int_{-\infty}^{\infty} p(x) dx &= \int_{-\infty}^{\mu} ce^{-\frac{|x-\mu|}{b}} dx + \int_{\mu}^{\infty} ce^{-\frac{|x-\mu|}{b}} dx \\ &= \int_{-\infty}^{\mu} ce^{\frac{x-\mu}{b}} dx + \int_{\mu}^{\infty} ce^{-\frac{x-\mu}{b}} dx = cbe^{\frac{x-\mu}{b}} \Big|_{x=-\infty}^{\mu} - cbe^{-\frac{x-\mu}{b}} \Big|_{x=\mu}^{\infty} \\ &= 2bc = 1 \Rightarrow c = \frac{1}{2b}. \end{aligned}$$

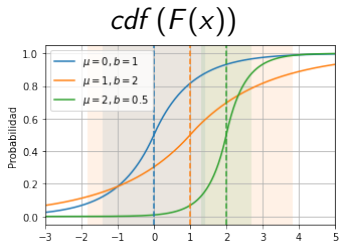
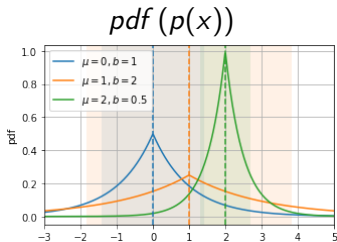
Entonces

$$p(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}, \text{ con } b > 0.$$

Distribución Laplace(μ, b)

Definición

$$p(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}, \text{ con } b > 0.$$



Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Laplace}(\mu, b)}[X] = \mu_X = \mu.$
- ▶ $\mathbb{V}_{X \sim \text{Laplace}(\mu, b)}[X] = \sigma_X^2 = 2b^2.$

¿Por qué usar distribución $\mathcal{N}(\mu, \sigma^2)$?

No acotada

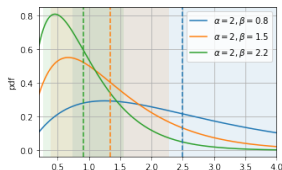
- ▶ ¿Y si los datos son solo positivos?
- ▶ ¿Y si los datos están acotados entre dos valores?

Distribución Gamma(α, β) = $\Gamma(\alpha, \beta)$

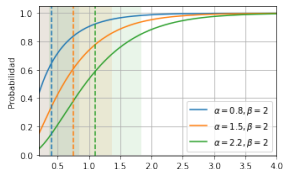
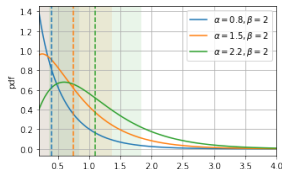
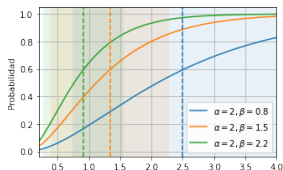
Definición

$$p(x) = \underbrace{\frac{\beta^\alpha}{\Gamma(\alpha)}}_{\text{normalizacion}} x^{\alpha-1} e^{-\beta x}, \text{ con } x \in (0, +\infty).$$

pdf ($p(x)$)



cdf ($F(x)$)



Distribución Gamma(α, β) = $\Gamma(\alpha, \beta)$

Definición

$$p(x) = \underbrace{\frac{\beta^\alpha}{\Gamma(\alpha)}}_{\text{normalizacion}} x^{\alpha-1} e^{-\beta x}, \text{ con } x \in (0, +\infty).$$

Nota

$\beta = 1/\mu$ promedio de ocurrencia de eventos.

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

$\frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-x/\beta}$
fase promedio de eventos.

Propiedades

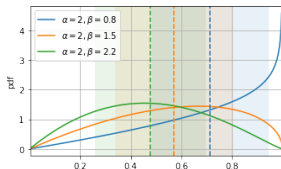
- ▶ $\mathbb{E}_{X \sim \Gamma(\alpha, \beta)}[X] = \mu_X = \frac{\alpha}{\beta}.$
- ▶ $\mathbb{V}_{X \sim \Gamma(\alpha, \beta)}[X] = \sigma_X^2 = \frac{\alpha}{\beta^2}.$

Distribución Beta(α, β)

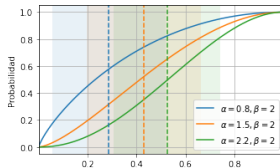
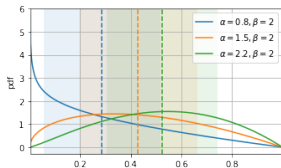
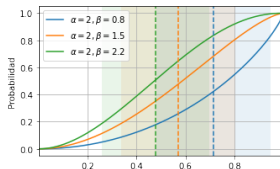
Definición

$$p(x) = \underbrace{\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}}_{\text{normalizacion}} x^{\alpha-1}(1-x)^{\beta-1}, \text{ con } x \in [0, 1].$$

pdf ($p(x)$)



cdf ($F(x)$)



Distribución Beta(α, β)

Definición

$$p(x) = \underbrace{\frac{1}{B(\alpha, \beta)}}_{\text{normalizacion}} x^{\alpha-1}(1-x)^{\beta-1}, \text{ con } x \in [0, 1].$$

Definición

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

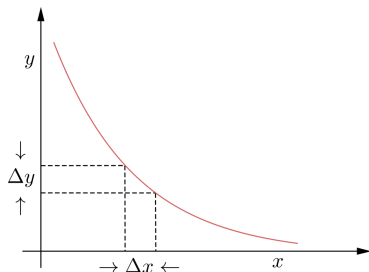
Propiedades

- ▶ $\mathbb{E}_{X \sim \text{Beta}(\alpha, \beta)}[X] = \mu_X = \frac{\alpha}{\alpha + \beta}.$
- ▶ $\mathbb{V}_{X \sim \text{Beta}(\alpha, \beta)}[X] = \sigma_X^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$

¿Cómo construir distribuciones?

Supuestos

- ▶ X variable aleatoria continua.
- ▶ $f : \mathbb{R} \rightarrow \mathbb{R}$ función invertible y monótona.
- ▶ $Y = f(X)$ nueva variable aleatoria.



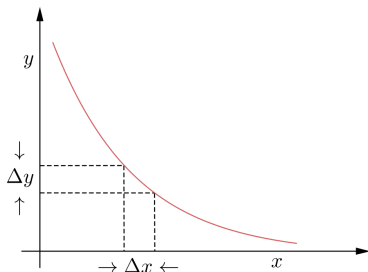
¿Cómo construir distribuciones?

Se tiene que

$$F_Y(y) = P(Y \leq y) = P(X \geq f^{-1}(y)) = 1 - F_X(f^{-1}(y)).$$

► Derivando con respecto a y :

$$p_Y(y) = -p_X(f^{-1}(y)) [f^{-1}]'(y) = -\frac{p_X(f^{-1}(y))}{f'(f^{-1}(y))}.$$



¿Cómo construir distribuciones?

Dos casos

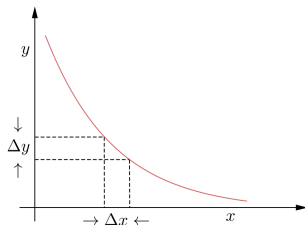
- Caso y monótona decreciente:

$$p_Y(y) = -\frac{p_X(f^{-1}(y))}{f'(f^{-1}(y))}.$$

- Caso y monótona creciente:

$$F_Y(y) = P(Y \leq y) = P(X \leq f^{-1}(y)) = F_X(f^{-1}(y))$$

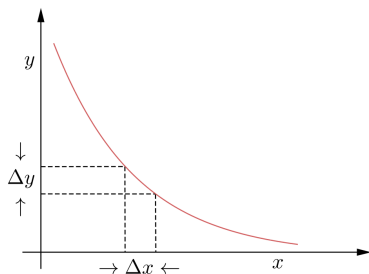
$$\Rightarrow p_Y(y) = \frac{p_X(f^{-1}(y))}{f'(f^{-1}(y))}.$$



¿Cómo construir distribuciones?

En general

$$p_Y(y) = \frac{p_X(f^{-1}(y))}{|f'(f^{-1}(y))|}.$$



Transformaciones

Ejemplo

- ▶ Si $X \sim \mathcal{N}(0, 1)$, ¿cómo se distribuye $Y = \mu + \sigma X$?

Desarrollo

- ▶ $f(x) = \mu + \sigma x$.
- ▶ $f'(x) = \sigma$.
- ▶ $f^{-1}(y) = \frac{y-\mu}{\sigma}$.
- ▶ Luego,

$$p_Y(y) = \frac{p_X(f^{-1}(y))}{|f'(f^{-1}(y))|} = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{y-\mu}{\sigma}\right)^2}{2}} \frac{1}{|\sigma|} = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$$

Conclusión

- ▶ $Y \sim \mathcal{N}(\mu, \sigma)$.

Transformaciones

Ejemplo

- Si $X \sim \mathcal{N}(\mu, \sigma)$, ¿cómo se distribuye $Y = e^X$?

Desarrollo

- $f(x) = e^x$.
- $f'(x) = e^x$.
- $f^{-1}(y) = \ln(y)$.
- Luego,

$$p_Y(y) = \frac{p_X(f^{-1}(y))}{|f'(f^{-1}(y))|} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}} \frac{1}{e^{\ln(y)}}$$
$$\Rightarrow p_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}}, \text{ con } y \in (0, +\infty).$$

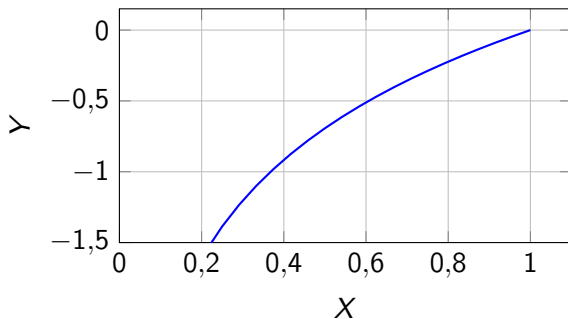
Conclusión

- ▶ Y sigue distribución llamada Log-normal ($Y \sim \text{Lognormal}(\mu, \sigma)$).

Transformaciones

Ejemplo

- Si $X \sim U(0, 1)$, ¿cómo se distribuye $Y = \ln(X)$?



Transformaciones

Ejemplo

- Si $X \sim U(0, 1)$, ¿cómo se distribuye $Y = \ln(X)$?

Desarrollo

- $f(x) = \ln(x)$.
- $f'(x) = \frac{1}{x}$.
- $f^{-1}(y) = e^y$.
- Luego,

$$p_Y(y) = \frac{p_X(f^{-1}(y))}{|f'(f^{-1}(y))|} = 1 \frac{1}{\frac{1}{e^y}} = e^y, \text{ con } y \in (-\infty, 0].$$

Transformaciones – ¿y si f no es monótona?

Por ejemplo

- ▶ $X \sim \mathcal{N}(0, 1)$.
- ▶ $Y = X^2$.

Separar parte negativa y positiva

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\ \Rightarrow p_Y(y) &= \frac{p_X(\sqrt{y})}{2\sqrt{y}} - \frac{p_X(-\sqrt{y})}{-2\sqrt{y}} = \frac{1}{2\sqrt{y}}(p_X(\sqrt{y}) + p_X(-\sqrt{y})) \\ &\Rightarrow p_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \text{ con } y \in [0, +\infty). \end{aligned}$$

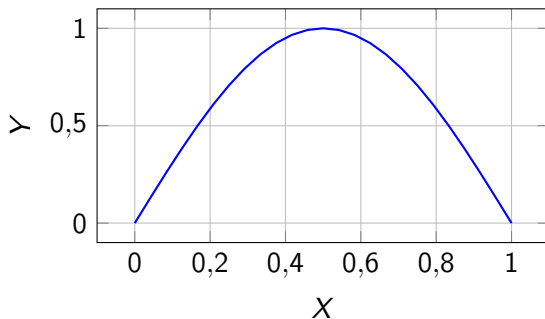
Conclusión

- ▶ Y sigue distribución llamada ji al cuadrado ($Y \sim \chi^2(1)$).

Transformaciones – ¿y si f no es monótona?

Ejemplo

- Si $X \sim U(0, 1)$, ¿cómo se distribuye $Y = \sin(\pi X)$?

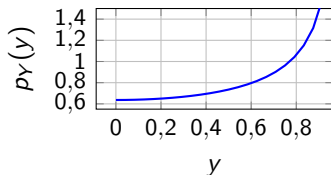


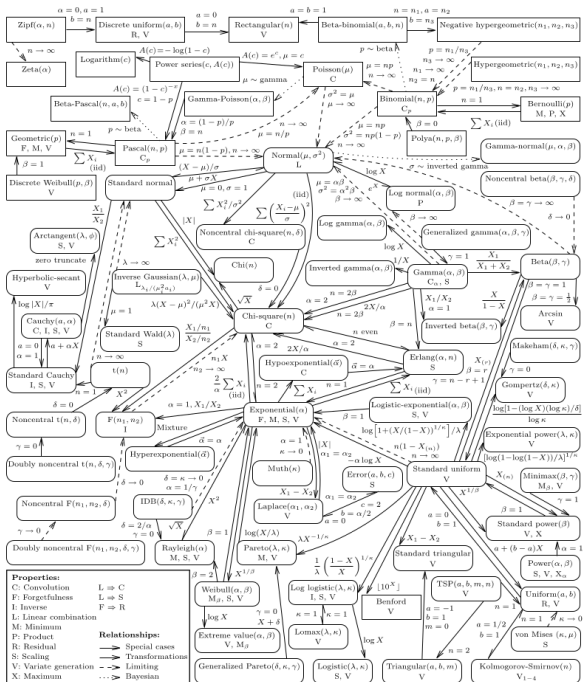
Transformaciones – ¿y si f no es monótona?

Desarrollo

- Si $X \sim U(0, 1)$, ¿cómo se distribuye $Y = \sin(\pi X)$?

$$\begin{aligned}F_Y(y) &= P(Y \leq y) = P(\sin(\pi X) \leq y) \\&= P(\pi X \leq \arcsin(y)) + P(\pi X > \pi - \arcsin(y)) \\&= P(\pi X \leq \arcsin(y)) + 1 - P(\pi X \leq \pi - \arcsin(y)) \\&= \frac{1}{\pi} \arcsin(y) + 1 - \frac{1}{\pi} (\pi - \arcsin(y)) = \frac{2}{\pi} \arcsin(y) \\&\Rightarrow p_Y(y) = \frac{2}{\pi \sqrt{1-y^2}}, \text{ con } y \in (0, 1)\end{aligned}$$





Extensión a varias variables aleatorias discretas

Probabilidad conjunta

$$P(X = x, Y = y) = f_{X,Y}(x, y).$$

$$P(X \leq x, Y \leq y) = F_{X,Y}(x, y) = \sum_{x'=-\infty}^x \sum_{y'=-\infty}^y f_{X,Y}(x', y').$$

		X				
	$f_{X,Y}(x, y)$	1	2	3	4	$f_Y(y)$
Y	1	0.1	0.33	0.15	0.04	0.62
	2	0	0.03	0.2	0.02	0.25
	3	0	0	0.05	0.08	0.13
	$f_X(x)$	0.1	0.36	0.4	0.14	

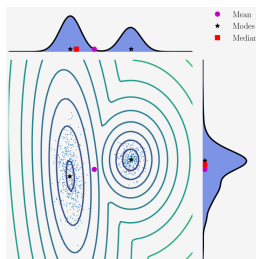
Probabilidad marginal

$$P(X = x) = f_X(x) = \sum_{y=-\infty}^{+\infty} f_{X,Y}(x, y) = \sum_{y=-\infty}^{+\infty} P(X = x, Y = y).$$

Extensión a varias variables aleatorias continuas

Probabilidad conjunta

$$P(X \leq x, Y \leq y) = F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y p_{X,Y}(x', y') dy' dx'.$$



Función de densidad de probabilidad (*pdf*) marginal

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y) dy.$$

Probabilidades y variables aleatorias

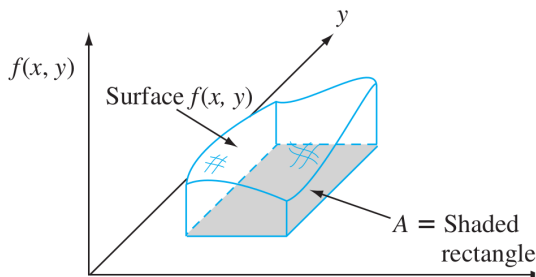
Eventos

- Caso discreto:

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

- Caso continuo:

$$P((X, Y) \in A) = \iint_{(x,y) \in A} p_{X,Y}(x, y) dx dy.$$



Probabilidades y variables aleatorias

Recuerdo de definiciones

- Probabilidad condicional:

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

- Independencia:

$$P(A, B) = P(A)P(B).$$

		X				
	$f_{X,Y}(x,y)$	1	2	3	4	$f_Y(y)$
Y	1	0.1	0.33	0.15	0.04	0.62
	2	0	0.03	0.2	0.02	0.25
	3	0	0	0.05	0.08	0.13
	$f_X(x)$	0.1	0.36	0.4	0.14	

Probabilidades y variables aleatorias

Y	$f_{Y X=3}(y)$
1	0.375
2	0.5
3	0.125

Extensión a varias variables aleatorias discretas

Probabilidad condicional

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

Independencia

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

Regla del producto

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

Teorema de Bayes

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

Extensión a varias variables aleatorias continuas

Probabilidad condicional

$$p_{X|Y}(x | y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

Independencia

$$p_{X,Y}(x, y) = p_X(x)p_Y(y).$$

Regla del producto

$$p_{X,Y}(x, y) = p_{X|Y}(x | y)p_Y(y).$$

Teorema de Bayes

$$p_{Y|X}(y | x) = \frac{p_{X|Y}(x | y)p_Y(y)}{p_X(x)}.$$

Extensión a varias variables aleatorias

		X				
	$f_{X,Y}(x,y)$	1	2	3	4	$f_Y(y)$
Y	1	0.1	0.33	0.15	0.04	0.62
	2	0	0.03	0.2	0.02	0.25
	3	0	0	0.05	0.08	0.13
$f_X(x)$		0.1	0.36	0.4	0.14	

Probabilidad marginal

$$P(X = x) = f_X(x) = \sum_{y=-\infty}^{+\infty} f_{X,Y}(x,y) = \sum_{y=-\infty}^{+\infty} f_{X|Y}(x | y)f_Y(y).$$

Función de densidad de probabilidad (pdf) marginal

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x,y)dy = \int_{-\infty}^{\infty} p_{X|Y}(x | y)p_Y(y)dy.$$

Extensión a varias variables aleatorias

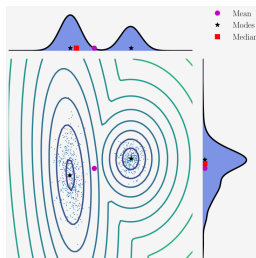
Valor esperado / esperanza

► Caso discreto:

$$\mathbb{E}_{X,Y}[X, Y] = \mathbb{E}_{X,Y}[(X, Y)] = \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} (x, y) f_{X,Y}(x, y).$$

► Caso continuo:

$$\mathbb{E}_{X,Y}[X, Y] = \mathbb{E}_{X,Y}[(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x, y) p_{X,Y}(x, y) dy dx.$$



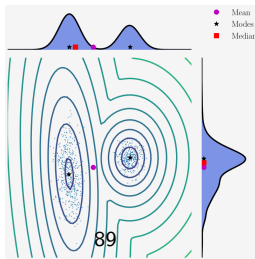
Probabilidad marginal... ¡De nuevo!

Probabilidad marginal

$$P(X = x) = f_X(x) = \sum_{y=-\infty}^{+\infty} f_{X|Y}(x | y) f_Y(y) = \mathbb{E}_Y[f_{X|Y}(x | y)].$$

Función de densidad de probabilidad (*pdf*) marginal

$$p_X(x) = \int_{-\infty}^{\infty} p_{X|Y}(x | y) p_Y(y) dy = \mathbb{E}_Y[p_{X|Y}(x | y)].$$

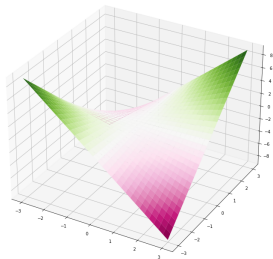


Extensión a varias variables aleatorias

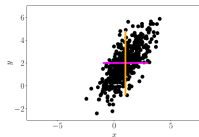
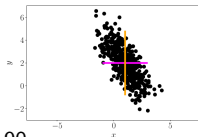
Covarianza y varianza

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}_{X,Y}[(X - \mathbb{E}_X[X])(Y - \mathbb{E}_Y[Y])] \\ &= \mathbb{E}_{X,Y}[XY] - \mathbb{E}_X[X]\mathbb{E}_Y[Y].\end{aligned}$$

$$\begin{aligned}\mathbb{V}[X] &= \text{Cov}(X, X) = \mathbb{E}_X[(X - \mathbb{E}_X[X])(X - \mathbb{E}_X[X])] \\ &= \mathbb{E}_X[X^2] - \mathbb{E}_X[X]^2.\end{aligned}$$



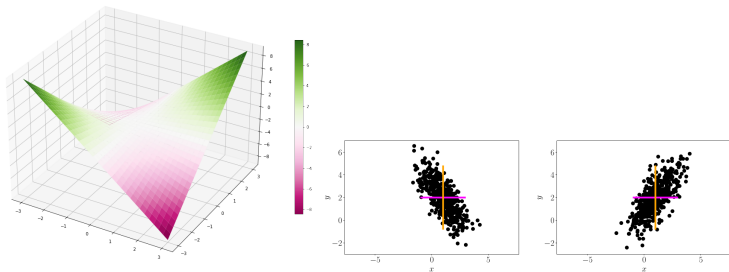
90



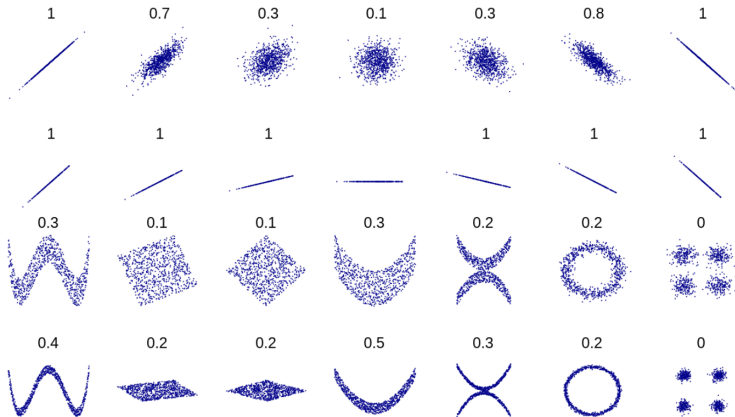
Extensión a varias variables aleatorias

Correlación

$$\text{corr}[X, Y] = \frac{\text{Cov}(X, Y)}{\sqrt{\text{V}[X]\text{V}[Y]}} \in [-1, 1].$$



Correlación



Correlación

Ejemplo

- ▶ X variable aleatoria con media 0 ($\mathbb{E}_X[X] = 0$) y $\mathbb{E}_X[X^3] = 0$.
- ▶ Sea $Y = X^2$.
- ▶ ¿ $\text{Cov}(X, Y)$?

Correlación

Ejemplo

- ▶ X variable aleatoria con media 0 ($\mathbb{E}_X[X] = 0$) y $\mathbb{E}_X[X^3] = 0$.
- ▶ Sea $Y = X^2$.
- ▶ ¿Cov(X, Y)?

Resultado

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[Y] = 0 - 0 = 0.$$

Extensión a varias variables aleatorias

Propiedades

- ▶ $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
- ▶ $\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y]$.
- ▶ $\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\text{Cov}(X, Y)$.
- ▶ $\mathbb{V}[X - Y] = \mathbb{V}[X] + \mathbb{V}[Y] - 2\text{Cov}(X, Y)$.

Demostración varianza de $X + Y$

$$\begin{aligned}\mathbb{V}[X + Y] &= \mathbb{E}\left[(X + Y - \mathbb{E}[X + Y])^2\right] \\ &= \mathbb{E}\left[(X - \mathbb{E}[X] + Y - \mathbb{E}[Y])^2\right]\end{aligned}$$

Extensión a varias variables aleatorias

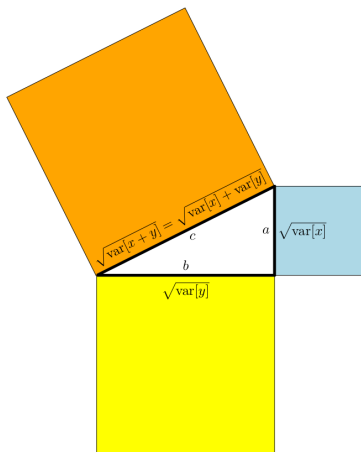
Demostración varianza de $X + Y$

$$\begin{aligned}\mathbb{V}[X + Y] &= \mathbb{E}\left[(X - \mathbb{E}[X])^2 + (Y - \mathbb{E}[Y])^2 + 2(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right] \\ &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] + \mathbb{E}\left[(Y - \mathbb{E}[Y])^2\right] + 2\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{V}[X] + \mathbb{V}[Y] + 2\text{Cov}(X, Y).\end{aligned}$$

Nota: producto interno de variables aleatorias

Dos variables no correlacionadas

$$\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y].$$



Distribuciones normales independientes

Supuesto

- ▶ Sean $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ e $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.
- ▶ Independientes: $p_{X,Y}(x, y) = p_X(x)p_Y(y)$.

Suma

- ▶ La variable aleatoria $X + Y$ es normal $\mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

En general

- ▶ La variable aleatoria $aX + bY$ sigue $\mathcal{N}(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$.

Distribuciones normales independientes

Ejemplo

- ▶ El precio de un litro de bencina en Santiago se determina según:

$$P = (1 + \text{IVA})P_{\text{enap}} + P_{\text{transporte}} + P_{\text{distribución}} + I_{\text{específico}},$$

- ▶ Donde:

- ▶ P_{enap} es el precio al que se vende en la refinería de Concón.
- ▶ $P_{\text{transporte}}$ es el precio de transportar hasta Maipú por el gasoducto.
- ▶ $P_{\text{distribución}}$ es un margen de venta para la distribuidora.
- ▶ IVA es el impuesto al valor agregado, actualmente 19 %.
- ▶ $I_{\text{específico}}$ es el impuesto específico, correspondiente a 6 UTM por m^3 .

Supuestos

- ▶ $P_{\text{enap}} \sim \mathcal{N}(500, 10000)$.
- ▶ $P_{\text{transporte}} \sim \mathcal{N}(10, 4)$.
- ▶ $P_{\text{distribución}} \sim \mathcal{N}(75, 400)$.⁹⁹

Distribuciones normales independientes

Ejemplo

$$P = (1 + \text{IVA})P_{\text{enap}} + P_{\text{transporte}} + P_{\text{distribución}} + I_{\text{específico}}.$$

- ▶ P es normal $\mathcal{N}(\mu_P, \sigma_P^2)$, con:
 - ▶ $\mu_P = (1 + 0,19)\mu_{P_{\text{enap}}} + \mu_{P_{\text{transporte}}} + \mu_{P_{\text{distribución}}} + 6 \cdot 52631 \frac{1}{1000}.$
 - ▶ $\sigma_P^2 = (1 + 0,19)^2 \sigma_{P_{\text{enap}}}^2 + \sigma_{P_{\text{transporte}}}^2 + \sigma_{P_{\text{distribución}}}^2.$
- ▶ P es normal $\mathcal{N}(\mu_P, \sigma_P^2) = \mathcal{N}(995,786, 14565).$

